



ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ
ΣΠΟΥΔΩΝ ΣΤΗ ΔΙΟΙΚΗΣΗ ΕΠΙΧΕΙΡΗΣΕΩΝ
MASTER IN BUSINESS ADMINISTRATION

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΣΤΗ ΔΙΟΙΚΗΣΗ ΕΠΙΧΕΙΡΗΣΕΩΝ

Διπλωματική Εργασία

ΠΡΟΒΛΕΨΗ ΠΩΛΗΣΕΩΝ ΜΕ GOOGLE TRENDS

της

ΕΛΕΝΗΣ ΠΑΠΑΔΟΠΟΥΛΟΥ

Υποβλήθηκε ως απαιτούμενο για την απόκτηση του μεταπτυχιακού
διπλώματος ειδίκευσης στη Διοίκηση Επιχειρήσεων
(με εξειδίκευση στη Χρηματοοικονομική Διοίκηση)

Δεκέμβριος 2018

Στην κόρη μου Αθανασία...

Ευχαριστίες

Για τη συγγραφή της παρούσας διπλωματικής εργασίας θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, κύριο Θεόδωρο Παναγιωτίδη για την καθοδήγηση, την υπομονή, τις επικοινωνιακές συμβουλές καθώς επίσης και για τον πολύτιμο χρόνο που μου διέθεσε, δίχως τα οποία δεν θα ήταν δυνατή η εκπόνηση αυτής της εργασίας.

Ευχαριστώ, επίσης, τον Καθηγητή κύριο Ηλία Κατσίκια και τον Καθηγητή κύριο Στυλιανό Φουντά του τμήματος Οικονομικών Επιστημών του Πανεπιστημίου για τη συμμετοχή τους ως μέλη της τριμελούς εξεταστικής επιτροπής.

Τέλος, θα ήθελα να εκφράσω την ευγνωμοσύνη μου στην οικογένεια μου και ιδιαιτέρως στην μητέρα μου, για την αμέριστη συμπαράσταση και βοήθεια που μου προσέφεραν σε όλη τη διάρκεια του μεταπτυχιακού.

Περίληψη

Η πρόβλεψη των πωλήσεων συμβάλλει σημαντικά στην απόδοση των επιχειρήσεων, κυρίως όσων έχουν υιοθετήσει μια περισσότερο πελατοκεντρική προσέγγιση. Οι εταιρείες αντιμετωπίζουν αρκετές προκλήσεις προκειμένου να προβλέψουν με ακρίβεια τις πωλήσεις, χωρίς να γνωρίζουν τη μελλοντική ζήτηση. Από το 2006, όμως, που εμφανίστηκε για πρώτη φορά η υπηρεσία Google Trends, η διαδικασία της πρόβλεψης πέρασε σε άλλη διάσταση, καθώς ένα μεγάλο πλεονέκτημα της συγκεκριμένης εφαρμογής είναι ότι συλλέγει “Big Data”, επεξεργάζεται τις πληροφορίες προς διευκόλυνση της ανάλυσης και στη συνέχεια, παρέχει τις πληροφορίες αυτές δωρεάν. Αυτή η έκρηξη δεδομένων προσέκλυσε αμέσως το ακαδημαϊκό ενδιαφέρον, παρακινώντας τους ταυτόχρονα να εξερευνήσουν αυτό το νέο τομέα που δείχνει να έχει πολλές προοπτικές. Σκοπός της παρούσας διπλωματικής εργασίας είναι η καταγραφή των διαφόρων μεθόδων που έχουν χρησιμοποιηθεί τα τελευταία χρόνια προκειμένου να προβλεφθούν οι πωλήσεις ή άλλοι παράγοντες που επηρεάζουν άμεσα είτε έμμεσα τις πωλήσεις μέσω των Google Trends. Για να αναλυθεί ο τρόπος με τον οποίο τα δεδομένα αναζήτησης ιστού συμβάλλουν στους διάφορους τομείς έρευνας, γίνεται αναφορά σε άλλες προηγούμενες μελέτες, κυρίως ερευνητικές εργασίες, που χρησιμοποίησαν την υπηρεσία Google Trends ή επηρεάστηκαν από αυτή.

Keywords: Google Trends, Forecasting, Search engine data, Sales

Περιεχόμενα

Ευχαριστίες.....	3
Περίληψη	4
1. Εισαγωγή.....	6
2. Ανασκόπηση Βιβλιογραφίας.....	8
2.1. Πίνακας Βιβλιογραφικής Ανασκόπησης	17
3. Google Trends.....	24
3.1. Συμπεριφορά καταναλωτή	24
3.2. Ιστορική αναδρομή	26
4. Πεδία Εφαρμογής.....	31
4.1. Πληροφοριακά συστήματα ή επιστήμη της πληροφορικής	31
4.2. Ιατρική και Βιοεπιστήμες.....	34
4.3. Οικονομία, Διοίκηση Επιχειρήσεων & Χρηματοοικονομικά	38
4.3.1. Ανεργία.....	38
4.3.2. Κατανάλωση.....	39
4.3.3. Χρηματιστήριο	40
4.3.4. Διοίκηση Επιχειρήσεων.....	42
4.3.5. Τουρισμός	43
4.3.6. Αγοροπωλησίες ακινήτων	45
4.4. Πολιτική Επιστήμη	46
5. Συμπεράσματα.....	48
Βιβλιογραφία	50

1. Εισαγωγή

Η πρόβλεψη των πωλήσεων συμβάλλει σημαντικά στην απόδοση των επιχειρήσεων, κυρίως όσων έχουν υιοθετήσει μια περισσότερο πελατοκεντρική προσέγγιση. Ανακριβείς προβλέψεις πιθανώς να οδηγήσουν τις επιχειρήσεις σε αδιάθετα αποθέματα και απώλειες, γεγονός που καθιστά τη διαδικασία της πρόβλεψης εξαιρετικά σημαντική. (Beheshti-Kashi, Karimi, Thoben, Lütjen, and Teucke, 2015).

Οι εταιρείες αντιμετωπίζουν αρκετές προκλήσεις προκειμένου να προβλέψουν με ακρίβεια τις πωλήσεις. Για παράδειγμα, θα πρέπει να προβλέψουν πόσα προϊόντα θα χρειαστεί να παράγουν χωρίς να γνωρίζουν τη μελλοντική ζήτηση. Ως εκ τούτου, η σωστή πρόβλεψη είναι απαραίτητη καθώς η παραγωγή του σωστού αριθμού προϊόντων είναι δύσκολο να επιτευχθεί διαφορετικά (Fissahn, 2001).

Ως σημείο εκκίνησης για την κατάρτιση των προβλέψεων παλαιότερα χρησιμοποιούσαν συνήθως τις πωλήσεις από προηγούμενα έτη. Επιπλέον, μπορεί να εξετάζαν τις ανεκτέλεστες παραγγελίες της εταιρείας, τα σχέδια μάρκετινγκ, την τιμολογιακή πολιτική ή κλαδικές μελέτες (Garrison and Noreen, 2005). Η επανάσταση, όμως, που έγινε την τελευταία δεκαετία, όσον αφορά τα δεδομένα αναζήτησης ιστού (search engine data), την εύκολη πρόσβαση και τη δωρεάν χρήση που παρέχουν, άλλαξε τελείως τον τρόπο με τον οποίο οι ερευνητές αναζητούν πλέον πληροφορίες και δεδομένα.

Αναλυτικότερα, αυτή η έκρηξη δεδομένων επέστησε την προσοχή των ερευνητών γύρω από τις χρήσεις και εφαρμογές των “Big Data” (Khoury and Ioannidis, 2014). Μία από τις σημαντικότερες εφαρμογές χρήσης των “Big Data” είναι η υποστήριξη διαφόρων οργανισμών και ομάδων που επιδιώκουν (1) να κατανοήσουν τις κοινωνικές αλλαγές & (2) να κάνουν προβλέψεις (Jun et al., 2018). Στη παρούσα διπλωματική εργασία, αναλύονται τα δεδομένα που προέρχονται από μηχανές αναζήτησης στο διαδίκτυο και ειδικότερα μέσω της εφαρμογής Google Trends, τα οποία αποτελούν ένα από τα παραδείγματα χρήσης και εφαρμογής των “Big Data”. Στόχος είναι η διερεύνηση του κατά πόσο μπορούν τα δεδομένα αυτά να χρησιμοποιηθούν στην πρόβλεψη των πωλήσεων.

Για να αναλυθεί ο τρόπος με τον οποίο τα δεδομένα αναζήτησης ιστού συμβάλλουν στους διάφορους τομείς έρευνας, γίνεται αναφορά σε άλλες προηγούμενες μελέτες που χρησιμοποίησαν την υπηρεσία Google Trends. Όσον αφορά την υπηρεσία Google Trends, εμφανίστηκε για πρώτη φορά στις 11 Μαΐου 2006 (Google, 2018). Ο ρόλος της είναι να παρέχει δεδομένα από αναζητήσεις που έχουν πραγματοποιηθεί στη μηχανή

αναζήτησης της Google, η οποία αποτελεί την κυρίαρχη μηχανή αναζήτησης καθώς το 2018 κατείχε το 92.37% του μεριδίου αγοράς παγκοσμίως (StatCounter, 2018).

Όπως γίνεται αντιληπτό, αυτή η ανακάλυψη άλλαξε τελείως την κατεύθυνση του ακαδημαϊκού ενδιαφέροντος, παρακινώντας την εξερεύνηση αυτού του νέου τομέα που δείχνει να έχει πολλές προοπτικές. Αξίζει να σημειωθεί ότι οι Ginsberg et al. (2009) και οι (Choi and Varian, 2009; Choi and Varian, 2012), που εργάζονται για την Google, ήταν οι πρώτοι που άνοιξαν τον ασκό του Αιόλου σχετικά με τις πολλές δυνατότητες πρόβλεψης που παρέχουν τα δεδομένα των Google Trends.

Στο πλαίσιο αυτό, σκοπός της παρούσας διπλωματικής εργασίας είναι η καταγραφή των διάφορων μεθόδων που έχουν χρησιμοποιηθεί τα τελευταία χρόνια προκειμένου να προβλεφθούν μέσω των Google Trends οι πωλήσεις, αλλά και άλλοι οικονομικοί ή μη παράγοντες που επηρεάζουν είτε άμεσα είτε έμμεσα τις πωλήσεις, καθώς επίσης και η καλύτερη κατανόηση των πεδίων εφαρμογής, στα οποία αναπτύσσεται η χρήση και εφαρμογή αυτών των δεδομένων. Για το σκοπό αυτό, από το σύνολο της βιβλιογραφίας, που έκαναν χρήση της υπηρεσίας Google Trends ή επηρεάστηκαν από αυτή, επιλέχθηκαν και αναλύθηκαν κυρίως ερευνητικές εργασίες.

Συνεπώς, στο επόμενο κεφάλαιο γίνεται μία βιβλιογραφική ανασκόπηση, όπου περιγράφονται οι σημαντικότερες ερευνητικές εργασίες, οι οποίες χρησιμοποίησαν δεδομένα αναζήτησης ιστού της εφαρμογής Google Trends. Στο τρίτο κεφάλαιο παρουσιάζεται ο τρόπος λειτουργίας της εφαρμογής Google Trends και περιγράφεται αναλυτικά η συμπεριφορά των καταναλωτών. Επίσης, απαντώνται ερωτήματα όπως αν η ηλεκτρονική αναζήτηση για ένα προϊόν συσχετίζεται με την αγορά αυτού του προϊόντος και άρα με τις πωλήσεις του. Στο τέταρτο κεφάλαιο αναλύονται τα πεδία εφαρμογής των Google Trends και τέλος, στο κεφάλαιο 5 συνοψίζονται τα βασικά συμπεράσματα που εξάγονται από την ανάλυση.

2. Ανασκόπηση Βιβλιογραφίας

Κατά τη διάρκεια των προηγούμενων ετών διεξάχθηκαν πολλές έρευνες, οι οποίες ασχολήθηκαν με τις δυνατότητες πρόβλεψης που παρέχουν τα δεδομένα που προέρχονται από μηχανές αναζήτησης. Μία από τις πρώτες σχετικές μελέτες, που αποτέλεσε το έναυσμα να ξεκινήσουν πολλοί ερευνητές να ενδιαφέρονται για τις δυνατότητες αξιοποίησης των Google Trends, ήταν αυτή των Ginsberg et al. (2009). Μέσω της μελέτης τους κατάφεραν να προβλέψουν τα τρέχοντα επίπεδα γρίπης χρησιμοποιώντας δεδομένα από σχετικά ερωτήματα που είχαν γίνει στη μηχανή αναζήτησης της Google. Για τους σκοπούς της έρευνας τους, κάθε ερώτημα αναζήτησης (search query) ήταν μία πλήρης ακολουθία όρων που προέρχονταν από ένα χρήστη. Δεν συνδύαζαν γλωσσικές παραλλαγές, συνώνυμα, μεταφράσεις ή λάθη στην ορθογραφία, ελπίζοντας να ασχοληθούν κάποια στιγμή στο μέλλον με αυτά. Η βάση δεδομένων που έφτιαξαν περιείχε 50 εκατομμύρια από τα πιο συχνά ερωτήματα αναζήτησης που σχετίζονται με το θέμα, ενώ δισεκατομμύρια ερωτήματα αναζήτησης που δεν εμφανίζονταν συχνά, εξαιρέθηκαν από την μελέτη τους. Για την εύρεση της τοποθεσίας κάθε ερωτήματος αναζήτησης χρησιμοποίησαν τη διεύθυνση ηλεκτρονικού πρωτοκόλλου (internet protocol address). Όσον αφορά τα δεδομένα, χρησιμοποίησαν όλες τις εβδομάδες από 28-9-2003 έως 11-3-2007, για τις οποίες τα Κέντρα Πρόληψης και Ελέγχου Νόσων (CDC) ανέφεραν μη μηδενικό ποσοστό ασθένειας παρόμοια με γρίπη (influenza like illness - ILI). Αυτό είχε ως αποτέλεσμα να έχουν 128 σημεία (κάθε εβδομάδα ήταν ένα σημείο) για να εκπαιδεύσουν το μοντέλο. Προστέθηκαν 42 επιπλέον εβδομάδες, από 18-3-2007 έως 11-5-2008, για την τελική αξιολόγηση του μοντέλου. Τελικά, κατάφεραν να δημιουργήσουν ένα μοντέλο που προέβλεπε με μεγάλη ακρίβεια τη δραστηριότητα των ιών της γρίπης σε πραγματικό χρόνο, σε μία έως δύο εβδομάδες νωρίτερα από την έκθεση που εκδίδουν τα CDC (Ginsberg et al., 2009).

Στη συνέχεια, οι Wu & Brynjolfsson, (2009) παρουσίασαν έναν πολύ καλό και ταυτόχρονα απλό τρόπο πρόβλεψης των μελλοντικών επιχειρηματικών δραστηριοτήτων όσον αφορά τον τομέα των ακινήτων, με τη χρήση δεδομένων που προέρχονται από μηχανές αναζήτησης. Εφαρμόζοντας τη μεθοδολογία τους για να προβλέψουν τις τάσεις των κατοικιών, διαπίστωσαν ότι ο δείκτης "House Price Index" (HPI) μπορούσε να προβλέψει με μεγάλη ακρίβεια τις μελλοντικές πωλήσεις και τιμές των κατοικιών. Κατά τη διάρκεια της περιόδου δειγματοληψίας, δηλαδή για την περίοδο 2007-2009, κάθε

ποσοστιαία αύξηση του δείκτη αναζήτησης κατοικιών (HPI) προέβλεπε 121.400 επιπλέον πωλήσεις σπιτιών για το επόμενο τρίμηνο (Wu and Brynjolfsson, 2009).

Από την άλλη πλευρά, οι Da et al. (2009) ήταν από τους πρώτους που χρησιμοποίησαν δεδομένα από την αναζήτηση στο διαδίκτυο στη χρηματοοικονομική επιστήμη. Αναλυτικότερα, πρότειναν μία νέα μέθοδο μέτρησης του ενδιαφέροντος των επενδυτών χρησιμοποιώντας το δείκτη “Google Search Volume Index” (SVI), ο οποίος σε ένα δείγμα 3000 μετοχών από τον Ιανουάριο του 2004 έως τον Ιούνιο του 2008, αποδείχθηκε πολύ ισχυρός καθώς (1) συσχετίζεται με το ενδιαφέρον των επενδυτών, (2) εντοπίζει πιο γρήγορα από ότι συνήθως το ενδιαφέρον των επενδυτών και (3) υπολογίζει ακόμη και τους μικροεπενδυτές. Όπως γίνεται αντιληπτό, ο όγκος αναζήτησης στο διαδίκτυο (internet search volume) είναι ένας αντικειμενικός τρόπος για να αποκαλυφθούν και να ποσοτικοποιηθούν τα συμφέροντα των επενδυτών και ως εκ τούτου είναι πολύ πιθανό να υπάρχουν και άλλες εφαρμογές που θα μπορούσαν να χρησιμοποιηθούν στα χρηματοοικονομικά (Da et al., 2009).

Την ίδια χρονιά, οι Askitas & Zimmermann (2009) ενδιαφέρθηκαν για τις πληροφορίες αναζήτησης που παρέχονται μέσω της Google, ως μέσο γρήγορης παρακολούθησης των οικονομικών κρίσεων. Χρησιμοποιώντας δεδομένα δοσμένα σε μηνιαία βάση από τον Ιανουάριο του 2004 έως τον Απρίλιο του 2009, απέδειξαν ότι υπάρχει μεγάλη συσχέτιση μεταξύ της αναζήτησης λέξεων-κλειδιών και των ποσοστών της ανεργίας στη Γερμανία. Το καινοτόμο στοιχείο που ανέδειξε η συγκεκριμένη έρευνα είναι ότι τα δεδομένα των Google Trends μπορούν να χρησιμοποιηθούν για να προβλέψουν την οικονομική συμπεριφορά, υπολογισμένα με παραδοσιακές στατιστικές μεθόδους (Askitas and Zimmermann, 2009). Την επόμενη χρονιά, οι D’Amuri & Marcucci (2010) πρότειναν τη χρήση ενός δείκτη αναζήτησης εργασίας, του Google Index (GI), ως τον καλύτερο δείκτη για την πρόβλεψη του ποσοστού ανεργίας των ΗΠΑ. Για να το καταφέρουν αυτό πραγματοποίησαν μία έρευνα εις βάθος, συγκρίνοντας και αναλύοντας πολλά μοντέλα που υιοθετούν τόσο τον προτιμώμενο δείκτη (GI), όσο και τους παραδοσιακούς δείκτες που αναφέρονται στις αιτήσεις ανέργων (initial claims) των Ηνωμένων Πολιτειών ή συνδυασμούς και των δύο. Διαπίστωσαν ότι τα μοντέλα που χρησιμοποίησαν συμπληρωματικά το δείκτη GI προβλέπουν με μεγαλύτερη ακρίβεια το μηνιαίο ποσοστό ανεργίας (D’Amuri and Marcucci, 2010).

Επιπλέον, οι Goel et al. (2010) προσπαθούν να προβλέψουν τη συμπεριφορά των καταναλωτών στο κοντινό μέλλον, ανάλογα με το τι τους ενδιαφέρει να αναζητούν στο διαδίκτυο. Πιο συγκεκριμένα, κατάφεραν να προβλέψουν τις εισπράξεις του box-office για ταινίες μεγάλου μήκους, τις πωλήσεις βιντεοπαιχνιδιών για τους πρώτους μήνες κυκλοφορίας τους και την κατάταξη των τραγουδιών στη λίστα «Billboard Hot 100». Διαπίστωσαν, επίσης, ότι η παρακολούθηση της αναζήτησης των καταναλωτών στο ίντερνετ γενικά μπορεί να ενισχύσει την απόδοση των βασικών μοντέλων, ειδικά όταν υπάρχει έλλειψη σε άλλες πηγές δεδομένων ή όταν η ακρίβεια της πρόβλεψης δεν μπορεί να βελτιωθεί σημαντικά με άλλο τρόπο (Goel et al., 2010).

Επιπλέον, οι Drake et al. (2011) προσπάθησαν να διερευνήσουν παράγοντες που επηρεάζουν τη ζήτηση πληροφοριών από μεριάς των επενδυτών γύρω από τις ανακοινώσεις κερδών και να παρέχουν πληροφορίες για το πώς η διακύμανση της ζήτησης πληροφοριών επηρεάζει την ανταπόκριση των κερδών της κεφαλαιαγοράς. Το διαδίκτυο είναι ένα εργαλείο μέσω του οποίου οι ειδήσεις φτάνουν στους επενδυτές και γι' αυτό οι συγγραφείς πιστεύουν ότι η ενημέρωση που παρέχεται μέσω των μηχανών αναζήτησης της Google είναι ένα μέσο που χρησιμοποιούν οι επενδυτές για να πάρουν πληροφορίες. Ως δεδομένα χρησιμοποίησαν τον ημερήσιο όγκο αναζήτησης της Google (daily Google search volume) για τα tickers 422 επιχειρήσεων από το S&P 500 για τα έτη 2005 έως 2008. Χρησιμοποίησαν αυτές τις εταιρίες επειδή είναι οι μεγαλύτερες και οικονομικά σημαντικότερες επιχειρήσεις στην οικονομία. Εξετάζοντας τα δεδομένα διαπίστωσαν ότι οι αναζητήσεις μέσω της Google αυξάνονται περίπου δύο εβδομάδες πριν από την ανακοίνωση των κερδών και συνεχίζουν σε υψηλά επίπεδα για μια περίοδο μετά την ανακοίνωση. Αυτό το εύρημα υποδηλώνει ότι η διάχυση των πληροφοριών δεν είναι στιγμιαία, με την ανακοίνωση της είδησης για κέρδη. Επίσης, διαπίστωσαν ότι η αναζήτηση πληροφοριών συσχετίζεται θετικά με τις ειδήσεις των μέσων μαζικής ενημέρωσης, ενώ συσχετίζεται αρνητικά με την αποστροφή της προσοχής των επενδυτών (Drake et al., 2011).

Οι Vosen & Schmidt (2011) σύγκριναν τον δείκτη καταναλωτικού κλίματος του Πανεπιστημίου Michigan (University of Michigan Consumer Sentiment Index – MCSI) με τον δείκτη εμπιστοσύνης των καταναλωτών (Consumer Confidence Index), χρησιμοποιώντας μηνιαία δεδομένα από την εφαρμογή Google Trends, για την περίοδο από τον Ιανουάριο του 2005 έως το Σεπτέμβριο του 2009. Τα αποτελέσματα έδειξαν ότι σχεδόν σε όλα τα πειράματα πρόβλεψης, ο δείκτης της Google ξεπέρασε τους δύο άλλους

δείκτες με τους οποίους συγκρίθηκε στη συγκεκριμένη έρευνα. Αυτό υποδηλώνει ότι η ενσωμάτωση πληροφοριών των Google Trends σε μία μελέτη μπορεί να προσφέρει σημαντικά οφέλη σε όσους ενδιαφέρονται να προβλέψουν την ιδιωτική κατανάλωση (Vosen & Schmidt, 2011).

Επίσης, μία πολύ σημαντική μελέτη που επικεντρώνεται στους οικονομικούς δείκτες, ήταν και αυτή των Choi & Varian (2012), οι οποίοι ανέφεραν ότι στόχος τους δεν είναι να προβλέψουν το μακρινό μέλλον με τη χρήση των Google Trends αλλά να προβλέψουν το παρόν (nowcasting). Πιο συγκεκριμένα, ασχολούνται με τις λιανικές πωλήσεις, τις πωλήσεις αυτοκινήτων και σπιτιών, τις αιτήσεις για χορήγηση επιδόματος ανεργίας και το σχεδιασμό τουριστικών προορισμών. Επιπλέον, δίνουν διευκρινήσεις και εξηγούν κάποιες απλές μεθόδους πρόβλεψης, θέλοντας να βοηθήσουν νέους ερευνητές να ξεκινήσουν τη δική τους ανάλυση (Choi and Varian, 2012). Σε συνέχεια της μελέτης των Choi and Varian (2012) οι Carrière-Swallow and Labbé (2013) προσπάθησαν με τη σειρά τους να προβλέψουν και αυτοί το παρόν (nowcasting). Πιο αναλυτικά, στην έρευνα τους εξετάζουν κατά πόσο η παρακολούθηση των συνηθειών περιήγησης των χρηστών στο διαδίκτυο μπορεί να αποκαλύψει κάτι σχετικά με τη συνολική συμπεριφορά των καταναλωτών σε μια αναδύομενη αγορά (emerging market). Χρησιμοποιώντας δεδομένα από ερωτήματα αναζήτησης (search queries) της Google, εισάγουν ένα δείκτη ο οποίος μετράει το ενδιαφέρον των καταναλωτών στη Χιλή για συναλλαγές αυτοκινήτων που πραγματοποιούνται μέσω του διαδικτύου και δοκιμάζουν αν τελικά αυτό βελτιώνει την αποτελεσματικότητα μοντέλων που προβλέπουν αντίστοιχα πωλήσεις αυτοκινήτων. Παρά τα σχετικά χαμηλά ποσοστά χρήσης του διαδικτύου, διαπίστωσαν ότι τα μοντέλα που ενσωματώνουν τον δείκτη αυτοκινήτου (Google Trends Automotive Index – GTAI) προσφέρουν σημαντικά οφέλη σχετικά με τους χρόνους παράδοσης πληροφοριών και μπορούν να εντοπίζουν καλύτερα τα σημεία καμπίς στα δεδομένα των πωλήσεων (Carrière-Swallow and Labbé, 2013).

Μια άλλη μελέτη που παρουσιάζει ιδιαίτερο ενδιαφέρον είναι αυτή των Geva et al. (2013). Σύμφωνα με τους συγγραφείς είναι πολύ μεγάλος ο αριθμός των ανθρώπων που χρησιμοποιούν δεδομένα από ιστότοπους κοινωνικών μέσων για την πρόβλεψη οικονομικών αποτελεσμάτων, όπως για παράδειγμα οι πωλήσεις. Ωστόσο, στην πράξη, τα δεδομένα αυτά είναι δαπανηρά τόσο για τη συλλογή όσο και για την επεξεργασία τους. Επιπλέον, οι προβλέψεις πωλήσεων που βασίζονται σε δεδομένα κοινωνικών μέσων ενδέχεται να παρεμποδιστούν ή να εμφανίσουν σημαντικές ανακρίβειες από την τάση

των ανθρώπων να περιορίσουν τα θέματα που συζητούν δημοσίως. Τα τελευταία χρόνια, έχει δημιουργηθεί μια νέα πηγή καταγραφής δεδομένων μέσω των μηχανών αναζήτησης. Είναι ενδιαφέρον ότι η σχέση μεταξύ αυτών των δύο σημαντικών πηγών δεδομένων δεν είχε μελετηθεί πριν από την συγκεκριμένη έρευνα, η οποία ήρθε για να καλύψει αυτό το κενό. Τα αποτελέσματα της έρευνας έδειξαν ότι τα μοντέλα πρόβλεψης που βασίζονται σε δεδομένα τάσεων (models based on search trend data) φαίνεται να αποδίδουν καλύτερα και να συμπληρώνουν τα μοντέλα που βασίζονται σε φόρουμ (forum-data-based models). Επίσης, οι δύο πηγές δεδομένων εμφανίζουν ουσιαστικά διαφορετικά πρότυπα όσον αφορά την ικανότητα πρόβλεψης (patterns of predictive capacity) με την πάροδο του χρόνου (Geva et al., 2013).

Το 2015, οι Fantazzini and Toktamysova (2015) θέλησαν και αυτοί να ασχοληθούν με την πρόβλεψη των πωλήσεων αυτοκινήτων, καθώς οι μακροπρόθεσμες προβλέψεις έχουν καίρια σημασία για την αυτοκινητοβιομηχανία λόγω της μεγάλης διάρκειας που απαιτείται για τις διαδικασίες ανάπτυξης και παραγωγής ενός νέου μοντέλου. Έχοντας αυτό κατά νου, η συγκεκριμένη έρευνα προτείνει νέα πολυπαραγοντικά μοντέλα (multivariate models) για την πρόβλεψη των πωλήσεων αυτοκινήτων, χρησιμοποιώντας οικονομικές μεταβλητές και δεδομένα που προήλθαν μέσω ερωτημάτων αναζήτησης στη μηχανή της Google. Τα δεδομένα που χρησιμοποιήθηκαν είναι οι μηνιαίες πωλήσεις από δέκα μάρκες αυτοκινήτων στη Γερμανία για την περίοδο από τον Ιανουάριο του 2001 έως και τον Ιούνιο του 2014, ενώ ο χρονικός ορίζοντας της πρόβλεψης είναι έως και 2 χρόνια μπροστά. Τα μοντέλα που συμπεριλάμβαναν τα δεδομένα αναζήτησης της Google στατιστικά υπερέβησαν τα αντίστοιχα ανταγωνιστικά μοντέλα για τις περισσότερες μάρκες αυτοκινήτων καθώς και τους χρονικούς ορίζοντες πρόβλεψης (Fantazzini and Toktamysova, 2015).

Την ίδια χρονιά ο Tetlock (2015) θέλησε να εξετάσει τη σχέση μεταξύ των χρηματοπιστωτικών αγορών και των μέσων ενημέρωσης (media), όπως για παράδειγμα τα δεδομένα αναζήτησης ιστού, τις αναρτήσεις σε κοινωνικά δίκτυα και τα άρθρα εφημερίδων. Για την ανάλυση του χρησιμοποίησε εφευρετικές εμπειρικές μεθόδους, κυρίως την ανάλυση κειμένου, προκειμένου να ποσοτικοποιήσει τις πολύτιμες πληροφορίες που κυκλοφορούν στα διάφορα μέσα σχετικά με τις σύγχρονες χρηματοπιστωτικές αγορές. Αυτά τα δεδομένα και αυτές οι μέθοδοι μπορούν να αποτελέσουν ισχυρά εργαλεία στα χέρια των επενδυτών καθώς καταφέρνουν να αντιμετωπίσουν μακροχρόνια προβλήματα των χρηματοπιστωτικών αγορών, όπως την

αστάθεια των τιμών των μετοχών που κυμαίνεται σε τόσο υψηλά επίπεδα. Γενικά, εστιάζει σε 3 σημεία: (1) οι πληροφορίες των μέσων ενημέρωσης μπορούν να ασκήσουν μεγάλη επιρροή στις χρηματοπιστωτικές αγορές, (2) ακόμη και όταν δεν υπάρχει κάποια ένδειξη, τα μέσα ενημέρωσης μπορούν να αποτελέσουν ένα χρήσιμο παράθυρο για τις πεποιθήσεις των επενδυτών και των μάνατζερ & (3) οι θεωρίες των μέσων ενημέρωσης και επικοινωνίας μπορούν να βοηθήσουν στην ανάπτυξη της εμπειρικής έρευνας (Tetlock, 2015).

Στην ίδια λογική με τους Choi and Varian (2012) και τους Carrière-Swallow and Labbé (2013), οι Youn and Cho (2016) παρουσιάζουν ένα μοντέλο πρόβλεψης (nowcasting) για τις πωλήσεις τηλεοράσεων χρησιμοποιώντας ως δεδομένα λέξεις-κλειδιά που αναζητήθηκαν στη μηχανή αναζήτησης της Google. Για την πρόβλεψη και αξιολόγηση των μεταβλητών χρησιμοποίησαν το «αυτοπαλίνδρομο - κινητού μέσου» μοντέλο ARIMA. Τα αποτελέσματα τους ήταν πολύ ικανοποιητικά και μπορούν να φανούν ιδιαίτερα χρήσιμα στις επιχειρήσεις που ασχολούνται με την κατασκευή ή εμπορία τηλεοράσεων (Youn and Cho, 2016). Επιπλέον, την ίδια χρονιά οι Sekhavat and Abdollahi (2016) διερεύνησαν αν η συμπεριφορά των καταναλωτών στο διαδίκτυο συσχετίζεται με την πραγματική τους συμπεριφορά, όσον αφορά την αγορά ηλεκτρονικών παιχνιδιών. Πιο συγκεκριμένα, στόχος της μελέτης τους ήταν να εξετάσουν σε ποιο βαθμό μπορούν να αξιοποιήσουν τα δεδομένα αναζήτησης μέσω της εφαρμογής Google Trends για να προβλέψουν την παγκόσμια σειρά κατάταξης ηλεκτρονικών παιχνιδιών σε κινητά, στο Ιράν. Βάσει των αποτελεσμάτων τους υπάρχει συσχέτιση μεταξύ των δεδομένων αναζήτησης ιστού και της κατάταξης των παιχνιδιών (Sekhavat and Abdollahi, 2016).

Μία άλλη σημαντική μελέτη με μεγάλη συνεισφορά στην επιστημονική κοινότητα είναι αυτή των Jun et al. (2017). Μέσω της συγκεκριμένης έρευνας παρουσιάζεται η δυνατότητα αξιοποίησης των μηχανών αναζήτησης υπό μία νέα διάσταση, αυτή της πρόβλεψης κατά αναλογία. Αναλυτικότερα, εισαγάγετε η αναλογική μέθοδος για την πρόβλεψη της ζήτησης, η οποία συγκρίνει τα δεδομένα των Google Trends και του Naver Trend και δημιουργεί αναλογίες. Για να το καταφέρουν αυτό έκαναν σύγκριση μεταξύ της εφαρμογής Google Trends και του Naver Trend. Μέσω του Naver Trend μπορεί κανείς να μελετήσει στατιστικά στοιχεία που προέρχονται από μία από τις κύριες μηχανές αναζήτησης που χρησιμοποιούν στην Κορέα. Όσον αφορά τα δεδομένα, η επισκεψιμότητα ιστού των ΗΠΑ εξετάστηκε για την περίοδο 2004-2017 μέσω των

Google Trends και της Κορέας για την περίοδο 2007-2017 μέσω του Naver Trend, σε εβδομαδιαία και μηνιαία βάση (Jun et al., 2017).

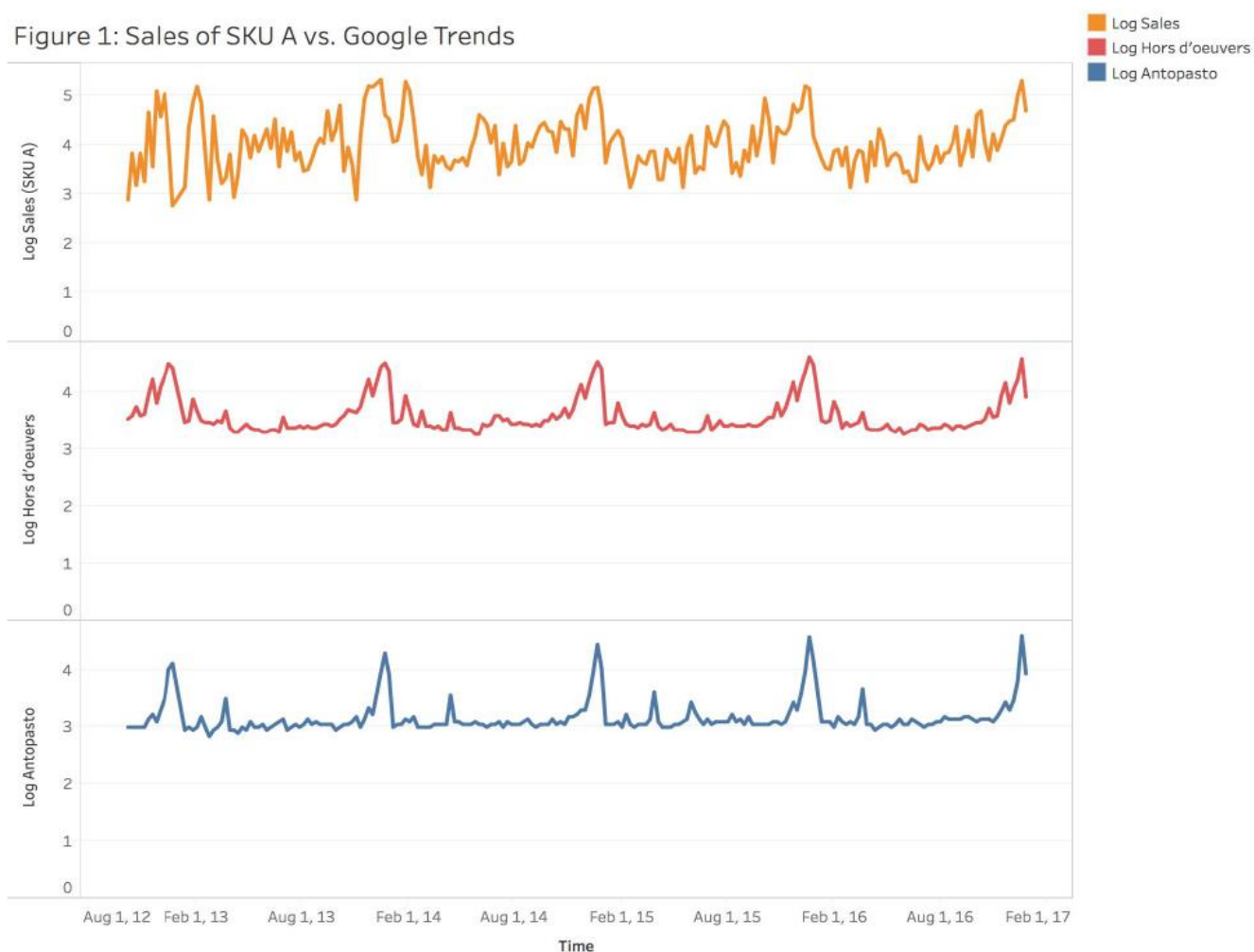
Την ίδια χρονιά, οι Elshendy et al. (2017) αναζήτησαν σημάδια οικονομικής ευαισθητοποίησης σχετικά με τα διαδικτυακά κοινωνικά μέσα και έλεγξαν την σημαντικότητα των δεδομένων τους στις οικονομικές προβλέψεις. Σε μία χρονική περίοδο 2 ετών, ανέλυσαν την τιμή του αργού πετρελαίου (West Texas Intermediate daily crude oil price) και τη σύγκριναν με τα διάφορα προγνωστικά δεδομένα που κυκλοφορούν από διάφορες πηγές, όπως Twitter, Google Trends, Wikipedia και τη βάση δεδομένων GDELT (Global Data on Events, Location and Tone). Για την πρόβλεψη και αξιολόγηση των μεταβλητών χρησιμοποιήθηκαν τα μοντέλα ARIMAX (Autoregressive Integrated Moving Average with Explanatory Variable). Τα αποτελέσματα έδειξαν ότι η συνδυασμένη ανάλυση των τεσσάρων μέσων πλατφόρμας παρέχει πολύτιμες πληροφορίες για τη διεξαγωγή οικονομικών προβλέψεων (Elshendy et al., 2017).

Καθώς οι μηχανές αναζήτησης αποτελούν κορυφαίο εργαλείο για τον προγραμματισμό διακοπών, οι Dergiades et al. (2018) προσπάθησαν να προβλέψουν τις αφίξεις των τουριστών στην Κύπρο. Χρησιμοποίησαν δεδομένα σε μηνιαία βάση από το 2004 έως το 2015 και έτρεξαν δύο ελέγχους μη αιτιότητας (non-causality testing). Αυτό το έκαναν γιατί όταν οι τουρίστες προέρχονται από χώρες με διαφορετικές γλώσσες και διαφορετικές πλατφόρμες μηχανών αναζήτησης, η πρόβλεψη των συνολικών διεθνών αφίξεων καθίσταται πολύ δύσκολη. Τελικά, παρατήρησαν ότι ο διορθωμένος δείκτης για διαφορετικές γλώσσες και διαφορετικές πλατφόρμες αναζήτησης, είναι καλύτερος σε σύγκριση με το μη προσαρμοσμένο. Η σπουδαιότητα της συγκεκριμένης έρευνας είναι μεγάλη, καθώς είναι η πρώτη έρευνα που καταφέρνει να βρει λύση στο πρόβλημα της γλώσσας και της πλατφόρμας αναζήτησης (Dergiades et al., 2018).

Τέλος, μία από τις πιο πρόσφατες μελέτες μέχρι σήμερα είναι αυτή των Boone et al. (2018) σκοπός της οποίας είναι να αποδείξει πως τα δεδομένα αναζήτησης ιστού, που δημιουργούνται από τους ίδιους τους χρήστες μέσω της εφαρμογής Google Trends, μπορούν να χρησιμοποιηθούν για τη μείωση των σφαλμάτων πρόβλεψης εκτός του δείγματος. Αναφέρουν ότι η διαδικασία ενσωμάτωσης των δεδομένων αναζήτησης ιστού στις πωλήσεις των επιχειρήσεων είναι σχετικά εύκολη υπόθεση αλλά μπορεί να αποτελέσει και μεγάλη πρόκληση. Οι λόγοι είναι ότι (1) δεν είναι προφανής ο τρόπος επιλογής των λέξεων-κλειδιών, καθώς η διαδικασία αυτή είναι ένα μείγμα δοκιμών,

σφαλμάτων αλλά και διαίσθησης, (2) απαιτείται ένα επίπεδο εμπιστοσύνης στα δεδομένα των τάσεων τουλάχιστον για βραχυπρόθεσμες προβλέψεις, χωρίς να υπάρχει σαφής γνώση του τρόπου με τον οποίο υπολογίζει το δείκτη η Google και (3) τα μοντέλα που προβλέπουν το παρόν (nowcasting) έχουν βραχυπρόθεσμους χρονικούς ορίζοντες πρόβλεψης ενώ οι εφοδιαστικές αλυσίδες συνήθως αντιμετωπίζουν μακροπρόθεσμα προβλήματα. Στο παρακάτω Σχήμα 1, με την κίτρινη γραμμή, αποτυπώνονται οι καταγεγραμμένες πωλήσεις ενός ειδικού φαγητού που αγοράστηκε και χρησιμοποιήθηκε ως ορεκτικό, με την κωδική ονομασία “SKU A”, ενώ με την κόκκινη και μπλε γραμμή αποτυπώνονται οι τιμές των δεικτών τάσης (trend indices) που προέκυψαν από την αναζήτηση των όρων “Hors d’oeuvres” και “Antipasto” αντίστοιχα. Οι όροι “Hors d’oeuvres” και “Antipasto” χρησιμοποιούνται συχνά για την περιγραφή του προϊόντος “SKU A”.

Figure 1: Sales of SKU A vs. Google Trends



Σχήμα 1: Πρόβλεψη πωλήσεων ενός προϊόντος μέσω των Google Trends

Πηγή: Boone et al. (2018). Can Google Trends Improve Your Sales Forecast?. Production and Operations Management

Τα δεδομένα που χρησιμοποίησαν οι Boone et al. (2018) προέρχονται από έναν εξειδικευμένο αντιπρόσωπο τροφίμων και μαγειρικών σκευών, ο οποίος τους έδωσε

πρόσβαση τόσο στις πωλήσεις όσο και στις τιμές πολλών προϊόντων. Επίσης, οι τάσεις βρίσκονται σε μία κλίμακα από το 1 έως το 100, όπου η τιμή 100 υποδηλώνει την κορυφαία δημοτικότητα ενός όρου αναζήτησης, ενώ η τιμή 50 δηλώνει ότι πρόκειται για το μισό της μέγιστης αξίας. Παρατηρώντας το διάγραμμα, γίνεται εύκολα αντιληπτό ότι οι πωλήσεις του προϊόντος “SKU A” ταιριάζουν πολύ με τη δημοτικότητα των δύο σχετικών όρων αναζήτησης. Για παράδειγμα, οι πωλήσεις για το “SKU A” κορυφώνονται γύρω από την ημέρα των Ευχαριστιών, περίπου την ίδια στιγμή που κορυφώνονται και οι αναζητήσεις για τους όρους «Antipasto» και «Hors d'oeuvres» (Boone et al., 2018).

2.1. Πίνακας Βιβλιογραφικής Ανασκόπησης

Στον παρακάτω πίνακα συνοψίζεται η βιβλιογραφική ανασκόπηση της παρούσας διπλωματικής εργασίας.

A/A	Τίτλος	Συγγραφείς	Μεθοδολογία	Δεδομένα	Θέμα
1	Detecting influenza epidemics using search engine query data.	Ginsberg et al. (2009)	Search query data, Google Trends, Epidemiology	Τα δεδομένα που χρησιμοποιήσαν ήταν δοσμένα σε εβδομαδιαία βάση μέσω της εφαρμογής Google Trends από 28-9-2003 έως 11-3-2007 (μόνο για τις εβδομάδες που τα CDC ανέφεραν μη μηδενικά ποσοστά ILI) και από 18-3-2007 έως 11-5-2008, καθώς επίσης, χρησιμοποίησαν δεδομένα από τα Κέντρα Πρόληψης και Ελέγχου Νόσων (CDC) των Ηνωμένων Πολιτειών.	Σκοπός της παρούσας μελέτης ήταν να καταφέρουν να προβλέψουν τα τρέχοντα επίπεδα γρίπης, χρησιμοποιώντας δεδομένα που δεν έχουν υποστεί καμία επεξεργασία, από σχετικά ερωτήματα που είχαν γίνει στη μηχανή αναζήτησης της Google. Κατάφεραν να δημιουργήσουν ένα σύστημα παρακολούθησης που προβλέπει τη δραστηριότητα των ιών της γρίπης με μεγάλη ακρίβεια σε πραγματικό χρόνο σε μία έως δύο εβδομάδες νωρίτερα από την έκθεση που εκδίδουν τα Κέντρα Πρόληψης και Ελέγχου Νόσων (CDC).
2	In Search of Attention.	Da et al. (2009)	Google Trends, Finance, forecasting, Investor attention	Τα δεδομένα που χρησιμοποίησαν σχετικά με τη συχνότητα των όρων αναζήτησης τα κατέβασαν από την εφαρμογή Google Trends, καθώς επίσης και τον εβδομαδιαίο δείκτη Search Volume Index (SVI) για τις μετοχές. Χρησιμοποίησαν, επίσης, μετοχές κυρίως από τον δείκτη Russell 3000 τις οποίες πήραν απευθείας από την εταιρία "Frank Russell and Company" για το χρονικό διάστημα: Ιανουάριος 2004-Ιούνιος 2008.	Η συγκεκριμένη έρευνα χρησιμοποιεί δεδομένα από την αναζήτηση στο διαδίκτυο στη χρηματοοικονομική επιστήμη. Αναλυτικότερα, προτείνει μία νέα μέθοδο μέτρησης του ενδιαφέροντος των επενδυτών χρησιμοποιώντας το δείκτη "Google Search Volume Index" (SVI), ο οποίος αποδείχθηκε πολύ ισχυρός καθώς επιτρέπει τον γρήγορο εντοπισμό του ενδιαφέροντος των επενδυτών, ακόμη και των μικροεπενδυτών. Όπως γίνεται αντιληπτό, ο όγκος αναζήτησης στο διαδίκτυο (internet search volume) είναι ένας αντικειμενικός τρόπος για να αποκαλυφθούν και να ποσοτικοποιηθούν τα συμφέροντα των επενδυτών.

A/A	Τίτλος	Συγγραφείς	Μεθοδολογία	Δεδομένα	Θέμα
3	The Future of Prediction: How Google Searches Foreshadow Housing Prices and Quantities.	Wu and Brynjolfsson (2009)	Online Search, Housing Index, Housing Trends, Predictions	Τα δεδομένα που χρησιμοποίησαν αφορούν την περίοδο 2007-2009 και προέρχονται από διάφορες πηγές: ο όγκος αναζητήσεων ερωτημάτων στο ίντερνετ (σε εβδομαδιαία βάση) προέρχεται από την εφαρμογή Google Trends, οι πωλήσεις σπιτιών από το Nation Association of Realtors για 50 κράτη των Ηνωμένων Πολιτειών και ο δείκτης "house price index" (HPI) από το Office of Federal Housing Enterprise Oversight.	Σκοπός της συγκεκριμένης μελέτης είναι η παρουσίαση ενός απλού τρόπου πρόβλεψης των μελλοντικών επιχειρηματικών δραστηριοτήτων στον τομέα των ακινήτων, με τη χρήση δεδομένων από μηχανές αναζήτησης. Εφαρμόζοντας τη μεθοδολογία τους για να προβλέψουν τις τάσεις των κατοικιών, διαπίστωσαν ότι ο δείκτης "Hause Price Index" (HPI) μπορούσε να προβλέψει με μεγάλη ακρίβεια τις μελλοντικές πωλήσεις και τιμές των κατοικιών.
4	Google Econometrics and Unemployment Forecasting.	Askitas and Zimmermann (2009)	Google, internet, keyword search, search engine, unemployment, predictions, time-series analysis	Τα δεδομένα όσον αφορά τα ποσοστά ανεργίας σε μηνιαία βάση προήλθαν από το Federal Employment Agency της Γερμανίας για το διάστημα: Ιανουάριος 2004 - Απρίλιος 2009, ενώ όσον αφορά τις χρονοσειρές αναζήτησης λέξεων-κλειδιών από την εφαρμογή Google Insights.	Το ενδιαφέρον, στην παρούσα έρευνα εστιάζεται στις πληροφορίες αναζήτησης που παρέχονται μέσω της Google ως μέσο γρήγορης παρακολούθησης των οικονομικών κρίσεων. Οι οικονομικές κρίσεις απαιτούν γρήγορες πληροφορίες προκειμένου να προβλεφθούν εγκαίρως, το οποίο είναι δύσκολο να συμβεί σε καιρούς διαρθρωτικών αλλαγών. Αποδείχθηκε μέσω της μελέτης ότι υπάρχει μεγάλη συσχέτιση μεταξύ της αναζήτησης των λέξεων-κλειδιών και των ποσοστών ανεργίας στη Γερμανία, υπολογισμένα με παραδοσιακές στατιστικές μεθόδους.
5	Google It! Forecasting the US Unemployment Rate with A Google Job Search Index.	D'Amuri and Marcucci (2010)	Google Econometrics, Forecast Comparison, Keyword search, US Unemployment, Time Series Models	Τα δεδομένα προήλθαν από διάφορες πηγές: Τα μηνιαία ποσοστά ανεργίας που εκδίδει το Bureau of Labor Statistics από το Current Employment Statistics, τα εβδομαδιαία ποσοστά ανεργίας από το Department of Labor των ΗΠΑ και τις μεταβλητές του δείκτη Google Index (GI) από την εφαρμογή Google Insights.	Στη συγκεκριμένη έρευνα εξετάζουν τη χρήση ενός δείκτη αναζήτησης εργασίας, του Google Index, ως τον καλύτερο δείκτη για την πρόβλεψη του ποσοστού ανεργίας των ΗΠΑ. Για να το καταφέρουν αυτό πραγματοποίησαν μία έρευνα εις βάθος, συγκρίνοντας και αναλύοντας πολλά μοντέλα που υιοθετούν τόσο τον προτιμώμενο δείκτη (GI), όσο και τους παραδοσιακούς δείκτες που αναφέρονται στις αιτήσεις ανέργων των Ηνωμένων Πολιτειών ή συνδυασμούς και των δύο. Διαπίστωσαν ότι τα μοντέλα που χρησιμοποιούν συμπληρωματικά το δείκτη GI προβλέπουν με μεγαλύτερη ακρίβεια το μηνιαίο ποσοστό ανεργίας.

A/A	Τίτλος	Συγγραφείς	Μεθοδολογία	Δεδομένα	Θέμα
6	Predicting consumer behavior with Web search.	Goel et al. (2010)	culture, predictions, Yahoo	Όσον αφορά τα δεδομένα: Τα έσοδα των ταινιών για τις ΗΠΑ προήλθαν από το IMDb για τη χρονική περίοδο Οκτώβριο 2008 - Σεπτέμβριο 2009, οι πωλήσεις και αξιολογήσεις των βιντεοπαιχνιδιών προήλθαν από VGChartz για τη χρονική περίοδο Σεπτέμβριο 2008 - Σεπτέμβριο 2009, τη σειρά κατάταξης των τραγουδιών από το "weekly Billboard rank" για τη χρονική περίοδο Μάρτιο 2009 - Σεπτέμβριο 2009 & τα δεδομένα αναζήτησης για τις ταινίες, τα βιντεοπαιχνίδια και τα τραγούδια προήλθαν από την μηχανή αναζήτησης του Yahoo.	Η συγκεκριμένη μελέτη ασχολήθηκε με την πρόβλεψη της συμπεριφοράς των καταναλωτών, ανάλογα με το τι τους ενδιαφέρει να αναζητούν στο διαδίκτυο. Πιο συγκεκριμένα, κατάφεραν να προβλέψουν τις εισπράξεις του box-office για ταινίες μεγάλου μήκους, τις πωλήσεις βιντεοπαιχνιδιών για τους πρώτους μήνες κυκλοφορίας τους και την κατάταξη των τραγουδιών στη λίστα «Billboard Hot 100». Διαπίστωσαν, επίσης, ότι η παρακολούθηση της αναζήτησης των καταναλωτών στο ίντερνετ γενικά μπορεί να ενισχύσει την απόδοση των βασικών μοντέλων, ειδικά όταν υπάρχει έλλειψη σε άλλες πηγές δεδομένων ή η ακρίβεια της πρόβλεψης δεν μπορεί να βελτιωθεί σημαντικά με άλλο τρόπο.
7	Investor Information Demand: Evidence from Google Searches around Earnings Announcements.	Drake et al. (2011)	Information Demand, Earnings Announcements, Information Content, Price Discovery	Χρησιμοποίησαν τον ημερήσιο όγκο αναζήτησης της Google για τα tickers των S&P 500 για τα έτη 2005 έως 2008 από την εφαρμογή Google Trend. Τις υπόλοιπες μεταβλητές που χρησιμοποίησαν τις πήραν από Compustat, Thomson SCD, Thomson 13F και τις βάσεις δεδομένων First Call.	Σκοπός της παρούσας έρευνας ήταν να διερευνήσουν παράγοντες που επηρεάζουν τη ζήτηση πληροφοριών των επενδυτών γύρω από τις ανακοινώσεις κερδών και να παρέχουν πληροφορίες για το πώς η διακύμανση της ζήτησης πληροφοριών επηρεάζει την ανταπόκριση των κερδών της κεφαλαιαγοράς. Διαπίστωσαν ότι οι αναζητήσεις της Google αυξάνονται περίπου δύο εβδομάδες πριν από την ανακοίνωση κερδών και συνεχίζουν σε υψηλά επίπεδα για μια περίοδο μετά την ανακοίνωση, υποδεικνύοντας ότι η διάχυση των πληροφοριών δεν είναι στιγμιαία. Επίσης, διαπίστωσαν ότι η αναζήτηση πληροφοριών συσχετίζεται θετικά με τις ειδήσεις των μέσων μαζικής ενημέρωσης και αρνητικά με την αποστροφή της προσοχής των επενδυτών.
8	Forecasting private consumption: survey-based indicators vs. Google trends.	Vosen and Schmidt (2011)	Google Trends, private consumption, forecasting, consumer sentiment indicators	Τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα μελέτη σχετικά με την ιδιωτική κατανάλωση προέρχονταν από την εφαρμογή Google Trends, ήταν μηνιαία και αφορούσαν την περίοδο από τον Ιανουάριο του 2005 έως το Σεπτέμβριο του 2009. Επιπλέον, η έρευνα στηρίχθηκε και σε δύο άλλους δείκτες: στο University of Michigan's Consumer Sentiment Index (MCSI) και στο Conference Board's Consumer Confidence Index (CCI).	Σκοπός της συγκεκριμένης μελέτης είναι η σύγκριση του δείκτη καταναλωτικού κλίματος του Πανεπιστημίου Michigan (MCSI) με το δείκτη εμπιστοσύνης των καταναλωτών (Consumer Confidence Index), ο οποίος υποδεικνύει τις τάσεις των Αμερικανών καταναλωτών. Τα αποτελέσματα έδειξαν ότι σχεδόν σε όλα τα πειράματα πρόβλεψης, ο δείκτης της Google ξεπέρασε τους δύο άλλους δείκτες με τους οποίους συγκρίθηκε. Αυτό υποδηλώνει ότι η ενσωμάτωση πληροφοριών των Google Trends σε μία μελέτη μπορεί να προσφέρει σημαντικά οφέλη στην πρόβλεψη της ιδιωτικής κατανάλωσης.

A/A	Τίτλος	Συγγραφείς	Μεθοδολογία	Δεδομένα	Θέμα
9	Predicting the Present with Google Trends.	Choi and Varian (2012)	Google Trends, Nowcasting, AR models, Fixed-effects models	Τα δεδομένα που χρησιμοποίησαν σε αυτό το άρθρο προέρχονται από διάφορες πηγές: τις πωλήσεις αυτοκινήτων, σπιτιών & τις λιανικές πωλήσεις τις πήραν από το U.S. Census Bureau, τον αριθμό των αιτήσεων για επιδόματα ανεργίας από το US Department of Labor και δεδομένα σχετικά με τουριστικούς προορισμούς από το Hong Kong Tourism Board, όλα σε μηνιαία βάση. Τα μηνιαία δεδομένα των ερωτημάτων αναζήτησης αναφορικά με τους παραπάνω τομείς τα κατέβασαν από την εφαρμογή Google Trends.	Σκοπός της παρούσας εργασίας είναι να παρουσιάσουν τον τρόπο χρήσης των δεδομένων που προέρχονται από τη μηχανή αναζήτησης της Google προκειμένου να προβλέψουν τις βραχυπρόθεσμες τιμές κάποιων βασικών οικονομικών δεικτών. Πιο συγκεκριμένα, ασχολούνται με τις πωλήσεις αυτοκινήτων και κατοικιών, τις αιτήσεις για χορήγηση επιδόματος ανεργίας και τον σχεδιασμό τουριστικών προορισμών.
10	Nowcasting with Google Trends in an Emerging Market.	Carrière-Swallow and Labbé (2013)	nowcasting, Google Trends, forecast accuracy, emerging markets	Ο δείκτης "Google Trends Automotive Index" (GTAI) που χρησιμοποίησαν στη συγκεκριμένη εργασία προήλθε από την εφαρμογή Google Trends μετά από την αναζήτηση λέξεων-κλειδιών σχετικά με τα αυτοκίνητα. Οι πωλήσεις αυτοκινήτων προήλθαν από την εθνική στατιστική υπηρεσία της Χιλής και ο δείκτης οικονομικής δραστηριότητας IMACEC από την κεντρική τράπεζα της Χιλής. Όλα τα δεδομένα ήταν δοσμένα σε μηνιαία βάση και αφορούσαν το διάστημα από τον Ιανουάριο του 2006 έως τον Μάιο του 2010.	Στην παρούσα έρευνα, εξετάζουν κατά πόσο η παρακολούθηση των συνηθειών περιήγησης των χρηστών στο διαδίκτυο μπορεί να αποκαλύψει κάτι σχετικά με τη συνολική συμπεριφορά των καταναλωτών σε μια αναδυόμενη αγορά (emerging market). Για το σκοπό αυτό, εισάγουν ένα δείκτη (GTAI), ο οποίος μετράει το ενδιαφέρον των καταναλωτών στη Χιλή για συναλλαγές αυτοκινήτων που πραγματοποιούνται μέσω του διαδικτύου και ελέγχουν αν μπορεί να βελτιώσει την αποτελεσματικότητα μοντέλων που προβλέπουν τις πωλήσεις αυτοκινήτων. Διαπίστωσαν ότι τα μοντέλα που ενσωματώνουν τον δείκτη (GTAI) μπορούν να εντοπίζουν καλύτερα τα σημεία καμπής στα δεδομένα των πωλήσεων.
11	Do Customers Speak Their Minds?	Geva et al. (2013)	Search trends, forums, word-of-mouth, consumers' interest, sales prediction, online data	Τα δεδομένα της συγκεκριμένης μελέτης είναι μηνιαία και αφορούν 20 μάρκες αυτοκινήτων που πουλιούνται στις Ηνωμένες Πολιτείες την τετραετία 2007-2010. Χρησιμοποιούν 3 διαφορετικές πηγές δεδομένων: (1) τα δεδομένα πωλήσεων τα άντλησαν από το site "www.autonews.com" (2) τα δεδομένα αναζήτησης από την εφαρμογή Google Trends και (3) τα δεδομένα των Forums από τη σάρωση του διαδικτύου μέσω της Google (Google's vast scan of the internet).	Όλο και περισσότεροι άνθρωποι χρησιμοποιούν δεδομένα από ιστότοπους κοινωνικών μέσων για την πρόβλεψη οικονομικών αποτελεσμάτων, όπως π.χ. των πωλήσεων. Ωστόσο, τα δεδομένα αυτά είναι δαπανηρά τόσο για τη συλλογή όσο και την επεξεργασία τους. Επιπλέον, οι προβλέψεις πωλήσεων που βασίζονται σε δεδομένα κοινωνικών μέσων ενδέχεται να εμφανίσουν ανακρίβειες λόγω της τάσης των ανθρώπων να περιορίσουν τα θέματα που συζητούν δημοσίως. Σκοπός της έρευνας είναι να συγκρίνει τις παραπάνω βάσεις δεδομένων με τις βάσεις δεδομένων αναζήτησης ιστού, όπως π.χ. η εφαρμογή Google Trends. Τα αποτελέσματα της έρευνας έδειξαν ότι τα μοντέλα πρόβλεψης που βασίζονται σε δεδομένα τάσεων υπερτερούν αλλά και συμπληρώνουν τα μοντέλα που βασίζονται σε φόρουμ.

A/A	Τίτλος	Συγγραφείς	Μεθοδολογία	Δεδομένα	Θέμα
12	Forecasting German car sales using Google data and multivariate models.	Fantazzini and Toktamysova (2015)	Car Sales, Forecasting, Google, Google Trends, Global Financial Crisis, Great Recession	Τα δεδομένα που χρησιμοποιήθηκαν όσον αφορά τις νέες καταχωρίσεις αυτοκινήτων στη Γερμανία προήλθαν από το Federal Motor Transport Authority (Kraftfahrt-Bundesamt), είναι σε μηνιαία βάση και αφορούν την περίοδο από τον Ιανουάριο του 2001 έως και τον Ιούνιο του 2014. Στο σύνολο πρόκειται για 162 παρατηρήσεις. Τα μηνιαία δεδομένα των ερωτημάτων αναζήτησης αναφορικά με τον τομέα της αυτοκινητοβιομηχανίας τα κατέβασαν από την εφαρμογή Google Trends και αφορούν 22 μάρκες αυτοκινήτων, διαθέσιμες από το 2001.	Στη συγκεκριμένη έρευνα οι συγγραφείς θέλησαν να προβλέψουν τις πωλήσεις αυτοκινήτων στη Γερμανία, με χρονικό ορίζοντα έως και 2 χρόνια μπροστά. Πρότειναν τη χρησιμοποίηση νέων πολυπαραγοντικών μοντέλων χρησιμοποιώντας οικονομικές μεταβλητές και δεδομένα που προήλθαν από ερωτήματα αναζήτησης στη Google. Τα αποτελέσματα, για τις περισσότερες μάρκες αυτοκινήτων, έδειξαν μια υπεροχή των μοντέλων που συμπεριλάμβαναν τα δεδομένα αναζήτησης της Google εν συγκρίσει με τα ανταγωνιστικά μοντέλα καθώς και τους χρονικούς ορίζοντες πρόβλεψης.
13	The Role of Media in Finance.	Tetlock (2015)	Media, Attention, Behavioral finance, Event studies, Textual analysis	Τα δεδομένα που χρησιμοποιήθηκαν στη συγκεκριμένη έρευνα σχετικά με τα μέσα ενημέρωσης προήλθαν από διάφορες πηγές δεδομένων όπως: άρθρα εφημερίδων, δεδομένα αναζήτησης ιστού και αναρτήσεις σε κοινωνικά δίκτυα. Όσον αφορά τις χρηματοπιστωτικές αγορές αναλύθηκαν στοιχεία με βάση την κατηγοριοποίηση του Dow Jones για την εταιρία Ravenpack και για τη χρονική περίοδο από το 2000 έως το 2010.	Στην παρούσα μελέτη εξετάζεται η σχέση μεταξύ των χρηματοπιστωτικών αγορών και των μέσων ενημέρωσης, όπως άρθρα εφημερίδων, δεδομένα αναζήτησης ιστού & αναρτήσεις σε κοινωνικά δίκτυα. Για την ανάλυση ο συγγραφέας χρησιμοποίησε εφευρετικές εμπειρικές μεθόδους κυρίως την ανάλυση κειμένου, προκειμένου να ποσοτικοποιήσει τις πολύτιμες πληροφορίες που κυκλοφορούν στα διάφορα μέσα σχετικά με τις σύγχρονες χρηματοπιστωτικές αγορές. Αυτά τα δεδομένα και αυτές οι μέθοδοι μπορούν να αποτελέσουν ισχυρά εργαλεία στα χέρια των επενδυτών και των μάνατζερ καθώς καταφέρνουν να αντιμετωπίσουν μακροχρόνια προβλήματα των χρηματοπιστωτικών αγορών, όπως την αστάθεια των τιμών των μετοχών.
14	Nowcast of TV market using Google Trend data.	Youn and Cho (2016)	ARIMA, Google trend, Nowcast, TV	Τα δεδομένα τους επικεντρώθηκαν γύρω από την έρευνα της λέξης "Τηλεόραση" ως λέξη-κλειδί στη μηχανή αναζήτησης της Google και αφορούσαν την χρονική περίοδο από το 2009 έως το 2013. Επίσης, χρησιμοποίησαν την ανάλυση δεδομένων της εφαρμογής Google Trends, για να κατατάξουν τα δεδομένα αναζήτησης με βάση τις χώρες και τις πόλεις.	Στη συγκεκριμένη έρευνα, οι συγγραφείς παρουσιάζουν ένα μοντέλο πρόβλεψης (nowcasting) για τις πωλήσεις τηλεοράσεων χρησιμοποιώντας ως δεδομένα λέξεις-κλειδιά που αναζητήθηκαν στη μηχανή αναζήτησης της Google. Για την πρόβλεψη και αξιολόγηση των μεταβλητών χρησιμοποίησαν το «αυτοπαλίνδρομο - κινητού μέσου» μοντέλο ARIMA. Τα αποτελέσματα τους ήταν πολύ ικανοποιητικά και μπορούν να φανούν ιδιαίτερα χρήσιμα στις επιχειρήσεις που ασχολούνται με την κατασκευή ή εμπορία τηλεοράσεων.

A/A	Τίτλος	Συγγραφείς	Μεθοδολογία	Δεδομένα	Θέμα
15	Can Google nowcast the market trend of Iranian mobile games?	Sekhavat and Abdollahi (2016)	Web search data, Google trends, mobile games, popularity nowcast, market trends	Τα δεδομένα που χρησιμοποιήθηκαν στη συγκεκριμένη μελέτη προέρχονται από δύο πηγές: τα δεδομένα αναζήτησης ιστού για παιχνίδια κινητών τα οποία άντλησαν από την εφαρμογή Google Trends και τα δεδομένα σχετικά με την κατάταξη των καλύτερων παιχνιδιών για κινητά τα οποία κατέβασαν από την "εφαρμογή Annie" (App Annie), μια εταιρία που πραγματοποιεί μελέτες για εφαρμογές παιχνιδιών.	Σκοπός της παρούσας μελέτης είναι η διερεύνηση σχετικά με τη συμπεριφορά των καταναλωτών στο διαδίκτυο και κατά πόσο αυτή συσχετίζεται με την πραγματική τους συμπεριφορά, όσον αφορά την αγορά ηλεκτρονικών παιχνιδιών. Πιο συγκεκριμένα, στόχος της μελέτης είναι να εξετάσουν σε ποιο βαθμό μπορούν να αξιοποιήσουν τα δεδομένα αναζήτησης μέσω της εφαρμογής Google Trends για να προβλέψουν την παγκόσμια σειρά κατάταξης ηλεκτρονικών παιχνιδιών σε κινητά, στο Ιράν. Βάσει των αποτελεσμάτων τους υπάρχει συσχέτιση μεταξύ των δεδομένων αναζήτησης ιστού και της κατάταξης των παιχνιδιών.
16	Forecasting by analogy using the web search traffic.	Jun et al. (2017)	Analogical forecasting, Web search traffic, Product life cycle, Search cycle, Google trends, Naver trend	Τα δεδομένα των μεταβλητών που χρησιμοποιήθηκαν προήλθαν από διάφορες πηγές: η επισκεψιμότητα ιστού των ΗΠΑ (2004-2017) και της Κορέας (2007-2017) προήλθαν από τα Google Trends και το Naver Trend αντίστοιχα σε εβδομαδιαία και μηνιαία βάση, οι πωλήσεις Netbook (2007-2013, ανά τρίμηνο) προήλθαν από το iSuppli (IDC), οι πωλήσεις Wii (2007-2014, ανά τρίμηνο) από Nintendo και οι πωλήσεις HEV & PHEV (U.S.) (2007-2017, ανά μήνα) από την ιστοσελίδα Hybridcars.com.	Σκοπός της παρούσας έρευνας είναι η δυνατότητα αξιοποίησης των μηχανών αναζήτησης υπό μία νέα διάσταση, αυτή της πρόβλεψης κατά αναλογία. Για να το καταφέρουν αυτό έκαναν σύγκριση μεταξύ της εφαρμογής Google Trends και του Naver Trend. Μέσω του Naver Trend μπορεί κανείς να μελετήσει στατιστικά στοιχεία που προέρχονται από μία από τις κύριες μηχανές αναζήτησης που χρησιμοποιούν στην Κορέα. Αναλυτικότερα, εισάγετε η αναλογική μέθοδος για την πρόβλεψη της ζήτησης, η οποία συγκρίνει τα δεδομένα των Google Trends και του Naver Trend και δημιουργεί αναλογίες.
17	Using four different online media sources to forecast the crude oil price.	Elshendy et al. (2017)	Financial forecast, Global Data on Events, Location and Tone, Google Trends, Oil price, Twitter, Wikipedia, ARIMAX	Τα δεδομένα που χρησιμοποιήθηκαν στη συγκεκριμένη μελέτη αφορούσαν τις τιμές αργού πετρελαίου (West Texas Intermediate daily crude oil price) και προήλθαν από 4 πλατφόρμες: (1) Twitter, (2) Google Trends, (3) Wikipedia & (4) τη βάση δεδομένων GDELT (Global Data on Events, Location and Tone). Από τα δεδομένα που ήταν εβδομαδιαία αφαιρέθηκαν τα Σαββατοκύριακα και οι αργίες, δηλαδή 5 ημέρες/εβδομάδα ήταν μία παρατήρηση. Στοιχεία από 1 Απριλίου 2013 έως 14 Μαρτίου 2015 χρησιμοποιήθηκαν για να εκπαιδεύσουν το μοντέλο ενώ τα στοιχεία από 15 Μαρτίου 2015 έως 1 Απριλίου 2015 χρησιμοποιήθηκαν για τις προβλέψεις.	Στη παρούσα εργασία, οι συγγραφείς αναζήτησαν σημάδια οικονομικής ευαισθητοποίησης σχετικά με τα διαδικτυακά κοινωνικά μέσα και έλεγξαν την σημαντικότητα των δεδομένων τους στις οικονομικές προβλέψεις. Σε μία χρονική περίοδο 2 ετών, ανέλυσαν την τιμή του αργού πετρελαίου και τη σύγκριναν με τα διάφορα προγνωστικά δεδομένα που κυκλοφορούν από διάφορες πηγές, όπως Twitter, Google Trends, Wikipedia και τη βάση δεδομένων GDELT. Για την πρόβλεψη και αξιολόγηση των μεταβλητών χρησιμοποιήθηκαν τα μοντέλα ARIMAX. Τα αποτελέσματα έδειξαν ότι η συνδυαστική ανάλυση των παραπάνω μέσων παρέχει πολύτιμες πληροφορίες στη διεξαγωγή οικονομικών προβλέψεων.

A/A	Τίτλος	Συγγραφείς	Μεθοδολογία	Δεδομένα	Θέμα
18	Google Trends and tourists' arrivals: Emerging biases and proposed corrections.	Dergiades et al. (2018)	Web search intensity, Google Trends, Tourists' arrivals	Όσον αφορά τα δεδομένα: οι μηνιαίες χρονολογικές σειρές δεδομένων για τις αφίξεις τουριστών στην Κύπρο μαζί με τα δεδομένα αναζήτησης για τις κατάλληλα επιλεγμένες λέξεις-κλειδιά, το εύρος των οποίων είναι από τον Ιανουάριο του 2004 έως τον Απρίλιο του 2016 (148 παρατηρήσεις) προήλθαν από την εφαρμογή Google Trends, ενώ οι συνολικές αφίξεις τουριστών στην Κύπρο προήλθαν από τη Στατιστική Υπηρεσία της Κύπρου (είναι διαθέσιμες μέχρι τον Δεκέμβριο του 2015, 144 παρατηρήσεις).	Στην παρούσα έρευνα προσπάθησαν να προβλέψουν τις αφίξεις των τουριστών στην Κύπρο. Χρησιμοποίησαν δεδομένα σε μηνιαία βάση από το 2004 έως το 2015 και έτρεξαν δύο ελέγχους μη αιτιότητας (non-causality testing). Αυτό το έκαναν γιατί όταν οι τουρίστες προέρχονται από χώρες με διαφορετικές γλώσσες και διαφορετικές πλατφόρμες μηχανών αναζήτησης, η πρόβλεψη των συνολικών διεθνών αφίξεων καθίσταται πολύ δύσκολη. Τελικά, παρατήρησαν ότι ο διορθωμένος δείκτης για διαφορετικές γλώσσες και διαφορετικές πλατφόρμες αναζήτησης, είναι καλύτερος σε σύγκριση με το μη προσαρμοσμένο.
19	Can Google Trends Improve Your Sales Forecast?	Boone et al. (2018)	forecasting, search queries, big data	Τα δεδομένα που χρησιμοποίησαν στη συγκεκριμένη μελέτη προήλθαν από έναν εξειδικευμένο αντιπρόσωπο τροφίμων και μαγειρικών σκευών, ο οποίος τους έδωσε πρόσβαση τόσο στις πωλήσεις όσο και στις τιμές πολλών προϊόντων. Επιπλέον, τα δεδομένα των ερωτημάτων αναζήτησης αναφορικά με τον τομέα των τροφίμων τα κατέβασαν από την εφαρμογή Google Trends. Τα δεδομένα πριν από την 1η Ιανουαρίου 2016 τα χρησιμοποίησαν για την εκπαίδευση του μοντέλου, ενώ τα δεδομένα του 2016 τα χρησιμοποίησαν για να κάνουν την πρόβλεψη.	Σκοπός της έρευνας είναι να αποδείξει πως τα δεδομένα αναζήτησης ιστού (Google Trends), μπορούν να χρησιμοποιηθούν στη μείωση των σφαλμάτων πρόβλεψης εκτός δείγματος. Αναφέρουν ότι η διαδικασία ενσωμάτωσης αυτών των δεδομένων στις πωλήσεις των επιχειρήσεων μπορεί να αποτελέσει πρόκληση για τους εξής λόγους: (1) η επιλογή των λέξεων-κλειδιών είναι σύνθετη διαδικασία καθώς αποτελεί ένα μείγμα δοκιμών, σφαλμάτων & διαίσθησης, (2) απαιτείται εμπιστοσύνη στα δεδομένα των τάσεων τουλάχιστον για βραχυπρόθεσμες προβλέψεις, χωρίς να υπάρχει σαφής γνώση του τρόπου με τον οποίο υπολογίζει το δείκτη η Google και (3) τα μοντέλα που προβλέπουν το παρόν έχουν βραχυπρόθεσμους ορίζοντες πρόβλεψης ενώ οι εφοδιαστικές αλυσίδες συνήθως αντιμετωπίζουν μακροπρόθεσμα προβλήματα.

3. Google Trends

Υπάρχουν διάφοροι λόγοι που εξηγούν γιατί τα Google Trends αποτελούν μια τόσο δημοφιλή πηγή αναζήτησης δεδομένων (Jun et al., 2018). Ένας σημαντικός λόγος είναι ότι η αναζήτηση στο Google παρέχει στους καταναλωτές μια εξαιρετική πλατφόρμα στην οποία μπορούν να αναζητήσουν οτιδήποτε τους ενδιαφέρει. Με άλλα λόγια, αντικατοπτρίζει τις ανάγκες, τις επιθυμίες και τα ενδιαφέροντα των χρηστών. Ένας άλλος λόγος είναι ότι τα Google Trends είναι πολύ εύκολα στη χρήση και επιπλέον, η Google δεν συλλέγει απλά δεδομένα αλλά παρέχει και τη δυνατότητα σύγκρισης μεταξύ τους (Jun et al., 2014). Υπό το πρίσμα αυτό, οι Jun et al. (2018), διακρίνουν δύο βασικές θεωρίες όσον αφορά την αναζήτηση πληροφοριών στο ίντερνετ. Η πρώτη θεωρία έχει να κάνει με τα χαρακτηριστικά των αναζητήσεων όσον αφορά τη λήψη αποφάσεων από μέρους της διοίκησης ή του τμήματος μάρκετινγκ, ενώ η δεύτερη ασχολείται με την υιοθέτηση της τεχνολογίας σε τομείς που ασχολούνται με την οικονομία, την καινοτομία και γενικότερα την έρευνα και ανάπτυξη (Jun et al., 2018).

3.1. Συμπεριφορά καταναλωτή

Με την άνοδο των κινητών τηλεφώνων (smartphones), η προσβασιμότητα των καταναλωτών στο ίντερνετ έχει αυξηθεί κατακόρυφα σε σχέση με παλιά που έπρεπε κάποιος να έχει υπολογιστή για να μπορεί να συνδεθεί στο διαδίκτυο. Επομένως, οι καταναλωτές τα τελευταία χρόνια είναι πιο εξοικειωμένοι με την αναζήτηση στο ίντερνετ καθώς και με τις ηλεκτρονικές αγορές. Σχεδόν καθημερινά αναζητούν στο διαδίκτυο προϊόντα που τους ενδιαφέρουν, με βασικό στόχο να μεγιστοποιήσουν την ικανοποίησή τους, είτε αυτή είναι ψυχολογικής είτε οικονομικής φύσεως, αποφεύγοντας ταυτόχρονα το ρίσκο που εμπεριέχεται σε αυτή τη διαδικασία.

Σύμφωνα με τον Shim S. (2001), μια ηλεκτρονική αναζήτηση για ένα προϊόν συχνά καταλήγει και στην αγορά αυτού του προϊόντος. Υπάρχει, λοιπόν, μια θετική συσχέτιση μεταξύ της επιθυμίας κάποιου να κάνει μια αναζήτηση στο ίντερνετ και της αγοράς του προϊόντος. Οι To et al. (2007) ανέλυσαν πιο διεξοδικά τη θετική αυτή σχέση και εξέτασαν εις βάθος τα κίνητρα που ωθούν τελικά τον καταναλωτή να προχωρήσει σε μία αγορά. Πιο συγκεκριμένα, χωρίζουν τα κίνητρα σε δύο κατηγορίες: (1) τα κίνητρα που λειτουργούν με βάση το όφελος (utilitarian motivations) και επηρεάζονται από την ευκολία, την εξοικονόμηση κόστους, τη διαθεσιμότητα πληροφοριών και την επιλογή &

(2) τα κίνητρα που λειτουργούν με βάση το συναίσθημα (hedonic motivations) και επηρεάζονται από την περιπέτεια, την εξουσία και το κοινωνικό status. Αναφέρουν, επίσης, ότι τα κίνητρα που λειτουργούν με βάση το όφελος καταλήγουν πιο συχνά σε αγορά, ενώ τα κίνητρα που λειτουργούν με βάση το συναίσθημα έχουν άμεσο αντίκτυπο στην πρόθεση αναζήτησης και έμμεσο αντίκτυπο στην πρόθεση αγοράς (To et al., 2007).

Η συμπεριφορά των καταναλωτών που αναζητούν πληροφορίες στο διαδίκτυο επηρεάζεται συχνά από τα χαρακτηριστικά του προϊόντος αλλά και από τα χαρακτηριστικά του ίδιου του καταναλωτή που διεξάγει την έρευνα. Το μορφωτικό επίπεδο και οι γνώσεις των καταναλωτών διαδραματίζουν σημαντικό ρόλο στην όλη διαδικασία αναζήτησης και τελικά στη λήψη της απόφασης σχετικά με το αν θα αγοράσουν ή όχι κάποιο προϊόν (Raju et al., 1995). Γενικά, όσο βελτιώνονται οι γνώσεις των καταναλωτών, τόσο περισσότερες πληροφορίες αναζητούν γύρω από τα προϊόντα (Bettman and Park, 1980). Ωστόσο, ο συσχετισμός αυτός ισχύει μόνο μέχρι ένα συγκεκριμένο σημείο, δεδομένου ότι ο καταναλωτής με επαρκείς γνώσεις δεν θα χρειαστεί να πραγματοποιήσει πρόσθετες αναζητήσεις, με αποτέλεσμα τη μείωση της δραστηριότητας αναζήτησης από εκείνο το σημείο και έπειτα. Αν θέλαμε, λοιπόν, να αναπαραστήσουμε με γράφημα τη παραπάνω σχέση θα ήταν μια τετραγωνική καμπύλη αντίστροφη του σχήματος «U», με την παρακάτω εξίσωση να την περιγράφει:

$$y = \beta_0 - \beta_1 X + \beta_2 X^2$$

Στην περίπτωση αυτή, ένας καταναλωτής που βρίσκεται στη μέση της κορυφής κατέχει ένα ενδιάμεσο επίπεδο γνώσης και εμπλέκεται πιο ενεργά σε αναζητήσεις από τους άλλους καταναλωτές που κατέχουν είτε μεγαλύτερο είτε μικρότερο επίπεδο γνώσης (Rao and Sieben, (1992). Επιπλέον, ένας άλλος λόγος που οδηγεί τους καταναλωτές σε περισσότερες αναζητήσεις στο διαδίκτυο είναι όταν λείπουν πληροφορίες σχετικά με ένα συγκεκριμένο προϊόν, επειδή για παράδειγμα είναι καινούργιο στην αγορά. Ομοίως, οι καταναλωτές τείνουν να γίνονται πιο ενεργοί όταν συνδέουν ένα προϊόν με υψηλότερο επίπεδο κινδύνου, καθώς θέλουν να μάθουν περισσότερες πληροφορίες για τους κινδύνους που ελλοχεύουν (Beatty and Smith, 1987).

Σε κάθε περίπτωση, οι στατιστικές που κυκλοφορούν σχετικά με τις δραστηριότητες αναζήτησης στο διαδίκτυο αποτελούν καλές πηγές για παρακολούθηση, ανάλυση και πρόβλεψη, όσον αφορά την αποδοχή ενός προϊόντος ή όχι από τους καταναλωτές ή την κοινωνία. Λόγω αυτής της πιθανότητας, προτού η Google κυκλοφορήσει πλήρως τα δεδομένα αναζήτησης ιστού μέσω της εφαρμογής Google Trends και πριν από τη

δημοσίευση της έρευνας των Ginsberg et al. (2009), υπήρξαν ερευνητές που είχαν ήδη πραγματοποιήσει μελέτες χρησιμοποιώντας δεδομένα από άλλες μηχανές αναζήτησης, εκτός της Google (Jun et al., 2018). Πιο συγκεκριμένα, από το 2003 η Moe (2003) διερεύνησε τις δραστηριότητες των καταναλωτών σχετικά με αναζητήσεις πληροφοριών που πραγματοποιήσαν στο διαδίκτυο, και τις ταξινόμησε στις εξής δύο κατηγορίες: περιήγηση ή αναζήτηση, ανάλογα με τα κίνητρα και τα αποτελέσματά τους. Στη συνέχεια, οι Ettredge et al. (2005) υποστήριξαν ότι μπορούν να εντοπίσουν τις ανάγκες, τις επιθυμίες, τα ενδιαφέροντα και τις ανησυχίες των ανθρώπων μέσω δεδομένων που προέρχονται από μηχανές αναζήτησης (search engine data). Στο ίδιο επίπεδο έρευνας κινήθηκαν, επίσης, οι Zimmer et al. (2007) και οι Fenn and Raskino (2008).

Κατά αυτό τον τρόπο, συνεχίστηκε η έρευνα που επικεντρώνεται στις δραστηριότητες αναζήτησης (search activities) και καθώς η χρήση των μηχανών αναζήτησης αυξάνεται, οι εμπειρικές μελέτες που επιχειρήσαν να χρησιμοποιήσουν τα δεδομένα αυτά άρχισαν να πολλαπλασιάζονται. Η υπηρεσία Google Trends έγινε πλήρως διαθέσιμη το 2008 και έκτοτε υπήρξαν σημαντικές αλλαγές στη σχετική έρευνα (Jun et al., 2018).

3.2. Ιστορική αναδρομή

Η υπηρεσία Google Trends εμφανίστηκε για πρώτη φορά στις 11 Μαΐου 2006. Στη συνέχεια, η Google κυκλοφόρησε την εφαρμογή Google Insights στις 5 Αυγούστου 2008, μια πιο προηγμένη υπηρεσία που παρείχε στους χρήστες εξαιρετικά λεπτομερείς πληροφορίες σχετικά με τις τάσεις αναζήτησης. Στις 27 Σεπτεμβρίου 2012, η Google συνδυάζει τις δύο εφαρμογές Google Trends & Google Insights και προχωράει στη συγχώνευση τους (Google, 2018).

Ο ρόλος της εφαρμογής Google Trends είναι να παρέχει τα δεδομένα από τις αναζητήσεις που έχουν πραγματοποιηθεί στη μηχανή αναζήτησης της Google. Πόσο δημοφιλής είναι όμως η μηχανή αναζήτησης της Google; Η απάντηση είναι ότι η δημοτικότητα της είναι εκρηκτική, καθώς σύμφωνα με τον Sullivan (2016) κατέγραψε 1,2 τρισεκατομμύρια αναζητήσεις ετησίως το 2012 και εκτίμησε ότι οι αναζητήσεις θα φθάσουν τα δύο τρισεκατομμύρια το 2016. Αυτό μεταφράζεται σε 5,5 δισεκατομμύρια αναζητήσεις ημερησίως και 167 δισεκατομμύρια αναζητήσεις ανά μήνα, αφήνοντας πλέον ελάχιστες αμφιβολίες ότι έχουμε να κάνουμε με μια τυπικά μεγάλη πηγή δεδομένων. Επίσης, αξίζει να σημειωθεί ότι υπάρχουν κάποιοι περιορισμοί στη χρήση της μηχανής αναζήτησης της

Google. Για παράδειγμα, ορισμένες χώρες όπως η Κίνα, η Ιαπωνία, η Νότια Κορέα και η Ρωσία, δεν χρησιμοποιούν τη Google ως κυρίαρχη μηχανή αναζήτησης λόγω πολιτικών ή γλωσσικών ζητημάτων (Jun et al., 2018). Ωστόσο, τα παγκόσμια στατιστικά στοιχεία για το μερίδιο αγοράς των διαφόρων μηχανών αναζήτησης που κυκλοφορούν, όπως φαίνεται και από τον παρακάτω πίνακα, δείχνουν ότι σήμερα η Google κατέχει το 92.37% του μεριδίου αγοράς, παραμένοντας η κυρίαρχη μηχανή αναζήτησης με τρομερά μεγάλη διαφορά από τις υπόλοιπες (StatCounter, 2018).

Μηχανή Αναζήτησης	Μερίδιο αγοράς (Νοέμβριος 2018)
Google	92,37%
Bing	2,37%
Yahoo!	2,25%
Baidu	0,96%
YANDEX RU	0,63%
DuckDuckGo	0,36%

Ένας από τους λόγους που εκτοξεύθηκε η δημοτικότητα της Google ήταν η κυκλοφορία της εφαρμογής Google Insights, μιας δωρεάν υπηρεσίας που εισήχθη τον Αύγουστο του 2008. Πρόκειται για μια αναθεωρημένη έκδοση της εφαρμογής Google Trends, η οποία έγινε γρήγορα εξαιρετικά δημοφιλής όχι μόνο στους υπεύθυνους μάρκετινγκ αλλά και σε πολλούς άλλους ερευνητές. Η υπηρεσία επέτρεπε στους χρήστες να εντοπίζουν διάφορους όρους αναζήτησης και φράσεις που είχαν καταχωρηθεί στη μηχανή αναζήτησης της Google. Τώρα, με τη συγχώνευση των δύο εφαρμογών Google Trends & Google Insights, οι χρήστες μπορούν όχι μόνο να κατηγοριοποιούν και να οργανώνουν στατιστικά δεδομένα αναζήτησης ιστού αλλά και να τα ταξινομούν ανά γεωγραφικές περιοχές. Η υπηρεσία Google Trends παρέχει δεδομένα σχετικά με τον όγκο αναζήτησης (search volume) για συγκεκριμένους όρους αναζήτησης από το 2004 και έπειτα. Πιο συγκεκριμένα, τα αποτελέσματα αναζήτησης είναι ανάλογα με τον χρόνο και την τοποθεσία κάποιου ερωτήματος με βάση την παρακάτω διαδικασία:

- Κάθε σημείο δεδομένων διαιρείται με τον συνολικό αριθμό των αναζητήσεων του γεωγραφικού και χρονικού εύρους που αναπαριστά, προκειμένου να συγκριθεί η σχετική δημοτικότητα του. Διαφορετικά, οι περιοχές με τον μεγαλύτερο όγκο αναζητήσεων θα είχαν πάντα υψηλότερη κατάταξη.
- Οι αριθμοί που προκύπτουν υπολογίζονται σε μια κλίμακα εύρους από 0 έως 100, σύμφωνα με το ποσοστό κάποιου θέματος επί όλων των αναζητήσεων για όλα τα θέματα.
- Οι διαφορετικές περιοχές που εμφανίζουν το ίδιο ενδιαφέρον αναζήτησης για έναν όρο δεν έχουν πάντα τους ίδιους συνολικούς όγκους αναζητήσεων.

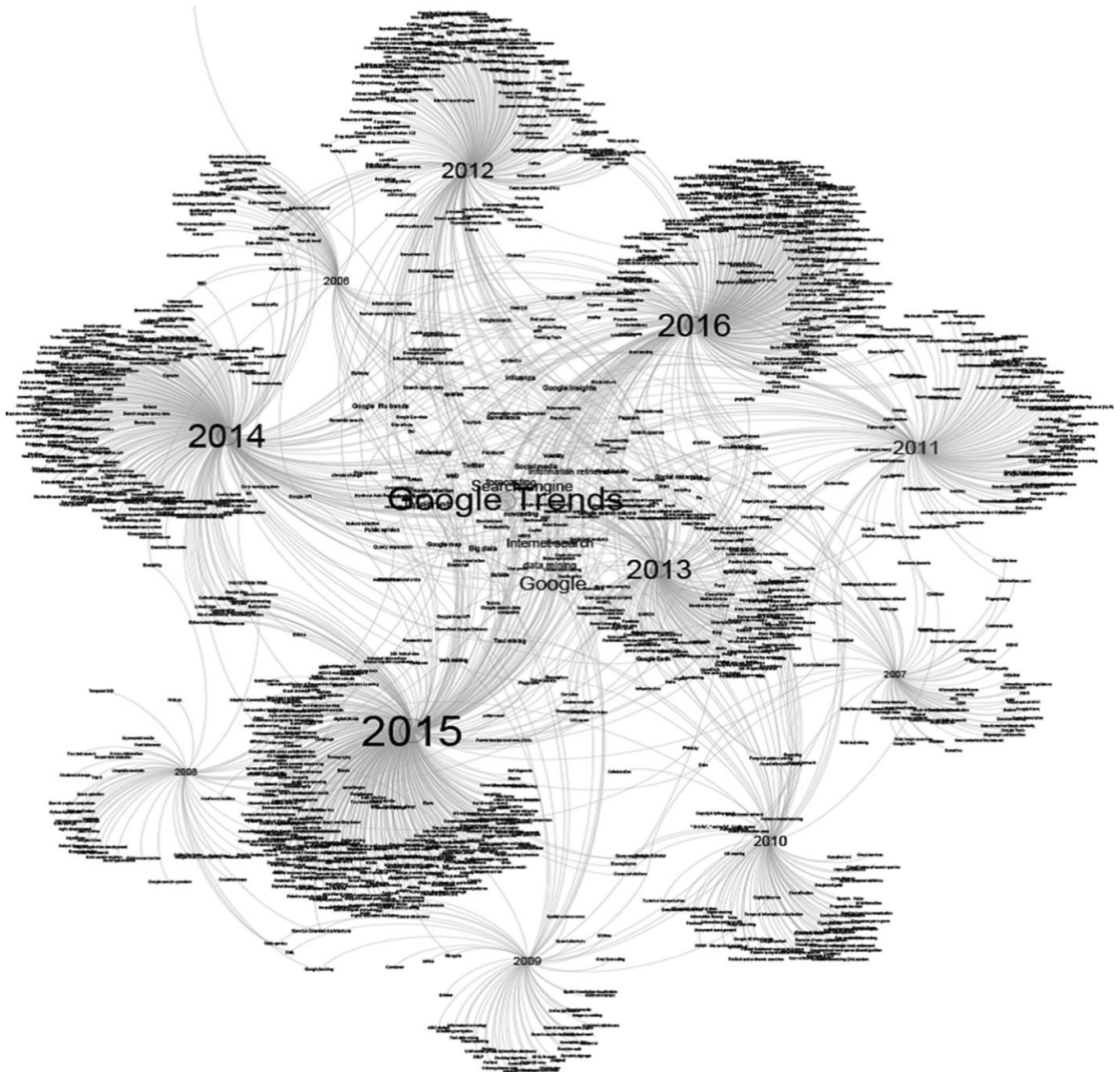
Ωστόσο, όπως εξηγήθηκε παραπάνω, ένας περιορισμός είναι ότι η Google παρέχει μόνο μια σχετική τιμή αναζήτησης και δεν παρέχει τον ακριβή όγκο αναζήτησης. Σε κάθε περίπτωση όμως, τα δεδομένα των Google Trends αποτελούν ένα αμερόληπτο δείγμα των δεδομένων αναζήτησης Google. Μόνο ένα ποσοστό των αναζητήσεων χρησιμοποιείται για τη συλλογή των δεδομένων της εφαρμογής Google Trends και υπάρχουν 2 τύποι δεδομένων:

- Τα δεδομένα σε πραγματικό χρόνο είναι ένα τυχαίο δείγμα αναζητήσεων από τις τελευταίες 7 ημέρες.
- Τα δεδομένα μη πραγματικού χρόνου, τα οποία είναι ένα τυχαίο δείγμα δεδομένων αναζήτησης Google που μπορεί να προέρχονται από το 2004 έως και 36 ώρες πριν από την αναζήτησή σας.

Μετά τη συλλογή, τα δεδομένα κατηγοριοποιούνται, συνδέονται σε κάποιο θέμα και αφαιρούνται τυχόν προσωπικά στοιχεία. Επιπλέον, από τα δεδομένα εξαιρούνται: (1) αναζητήσεις που πραγματοποιούνται από λίγα άτομα, (2) διπλότυπες αναζητήσεις, δηλαδή επαναλαμβανόμενες αναζητήσεις που έχουν πραγματοποιηθεί από το ίδιο άτομο μέσα σε σύντομο χρονικό διάστημα & (3) ειδικοί χαρακτήρες, όπως απόστροφοι, εισαγωγικά και άλλοι ειδικοί χαρακτήρες (Google, 2018).

Το Σχήμα 2 παρακάτω δείχνει τα αποτελέσματα μίας ανάλυσης δικτύου (network analysis) βάσει του έτους δημοσίευσης και των λέξεων-κλειδιών που επέλεξε ο εκάστοτε συγγραφέας, με εξαίρεση τις μείζονες λέξεις-κλειδιά που βρίσκονται στο κέντρο. Οι λέξεις-κλειδιά που εντοπίστηκαν στο κέντρο περιελάμβαναν τους όρους: Google Trends, μηχανή αναζήτησης, Google, διαδίκτυο, εξόρυξη δεδομένων (data mining), κοινωνικά

μέσα, επανεξέταση πληροφοριών, πρόβλεψη (forecasting), πρόβλεψη του παρόντος (nowcasting), μεγάλα δεδομένα (Big Data), γρίπη και Twitter. Παρατηρώντας το σχήμα, γίνεται αντιληπτό ότι κάθε χρονιά υπήρχαν πολλές μικρότερες λέξεις-κλειδιά σε αφθονία. Αυτές οι λέξεις αναζήτησης ήταν εκείνες που κατέλαβαν τις κορυφαίες βαθμολογίες όσον αφορά τη συχνότητα (Jun et al., 2018).



Σχήμα 2: Ανάλυση δικτύου βάσει λέξεων-κλειδιών & έτους δημοσίευσης

*Πηγή: Jun et al. (2018). Ten years of research change using Google Trends.
Technological Forecasting and Social Change*

Ένας άλλος λόγος που βοήθησε στην εκτόξευση της δημοτικότητας της εφαρμογής Google Trends ήταν άνθρωποι μέσα από την ίδια την εταιρία. Ο Ginsberg και οι

συνεργάτες του, που εργάζονται για την Google, έδειξαν ότι ασθένειες παρόμοιες με τη γρίπη μπορούν να ανιχνευθούν χρησιμοποιώντας την υπηρεσία Google Trends σε ερευνητικό έγγραφο που δημοσιεύτηκε στο περιοδικό “Nature”, υποκινώντας και άλλες μελέτες να ακολουθήσουν το δικό τους παράδειγμα (Ginsberg et al., 2009). Και παρόλο που υπήρξαν και άλλες μελέτες πριν από αυτή, η μελέτη των Ginsberg et al. (2009) αποτέλεσε σημείο αναφοράς για σχεδόν όλες τις μεταγενέστερες μελέτες που χρησιμοποίησαν την υπηρεσία Google Trends. Η συγκεκριμένη μελέτη παρουσιάζει τρία σημεία σχετικά με τη χρήση στατιστικών δεδομένων αναζήτησης: (1) η επισκεψιμότητα αναζήτησης (search traffic) που παρέχεται από την Google Trends αποτελεί ένα εξαιρετικό εργαλείο παρακολούθησης για την ταχεία ανίχνευση ενός συγκεκριμένου φαινομένου, (2) τα δεδομένα είναι αντιπροσωπευτικά, όπως αναφέραμε και παραπάνω, καθώς περισσότερες από 5,5 δισεκατομμύρια αναζητήσεις εκτελούνται καθημερινά μέσω της μηχανής αναζήτησης της Google & (3) τα δεδομένα αναζήτησης παρουσιάζουν μεγάλη συσχέτιση με τα κοινωνικά φαινόμενα, γεγονός που δείχνει τις δυνατότητες τους να εφαρμοστούν σε ένα ευρύ φάσμα τομέων. Στη συνέχεια, η Google πρότεινε άλλους τομείς, εκτός από την ιατρική, που μπορούν να επωφεληθούν από την υπηρεσία Google Trends. Οι Choi και Varian, που εργάζονται επίσης για την Google, υποστήριξαν ότι η εφαρμογή Google Trends μπορεί να χρησιμοποιηθεί για την πρόβλεψη του ποσοστού ανεργίας (Choi and Varian, 2009) καθώς και των πωλήσεων αυτοκινήτων ή σπιτιών (Choi and Varian, 2012). Τα άρθρα τους ανέφεραν ότι τα δεδομένα των Google Trends έχουν την ικανότητα να προβλέπουν το παρόν ή αλλιώς το βραχυπρόθεσμο μέλλον και μπορούν να χρησιμοποιηθούν ως ένα εξαιρετικό εργαλείο πρόβλεψης σε διάφορους τομείς (Jun et al., 2018).

Γίνετε φανερό ότι οι αρχικές μελέτες που έγιναν σχετικά με την υπηρεσία Google Trends, ήταν υποκινούμενες σε μεγάλο βαθμό από την Google και πραγματοποιήθηκαν από ακαδημαϊκούς επιστήμονες. Για το λόγο αυτό, έγιναν αντικείμενο μελέτης από πολλούς ερευνητές σε όλο τον κόσμο με διάφορα πεδία εφαρμογής. Στο επόμενο κεφάλαιο, αναλύονται οι τομείς στους οποίους οι μελέτες αυτές έχουν σημειώσει πρόοδο.

4. Πεδία Εφαρμογής

Διαβάζοντας τη σχετική βιβλιογραφία προκύπτουν τέσσερις (4) τομείς στους οποίους έχει διακριθεί κυρίως η έρευνα γύρω από τα Google Trends και την αναζήτηση μέσω λέξεων-κλειδιών:

- Πληροφοριακά συστήματα ή επιστήμη της πληροφορικής
- Ιατρική και Βιοεπιστήμες
- Οικονομία, Διοίκηση Επιχειρήσεων & Χρηματοοικονομικά
- Πολιτική επιστήμη

Οι Jun, Yoo & Choi (2018) δείχνουν στη μελέτη τους ότι οι τελευταίες έρευνες μέσω των Google Trends σχετίζονται περισσότερο με τα πληροφοριακά συστήματα και την επιστήμη της πληροφορικής. Πιο συγκεκριμένα, οι πιο πρόσφατες αναζητήσεις τείνουν να συγκεντρώνονται γύρω από τον όρο “Big Data”, έναν από τους δημοφιλέστερους όρους αναζήτησης τα τελευταία δέκα χρόνια, ο οποίος έχει σταδιακά επεκταθεί και στους τομείς της Ιατρικής και της Οικονομίας.

Συνεπώς, στα κεφάλαια που ακολουθούν αναλύονται τα σημαντικότερα αποτελέσματα που έχουμε αποκομίσει από επιστημονικές έρευνες που έχουν πραγματοποιηθεί στους παραπάνω τέσσερις τομείς, καθώς επίσης, χρήσεις και εφαρμογές σχετικά με τα “Big Data” και την εφαρμογή των Google Trends.

4.1. Πληροφοριακά συστήματα ή επιστήμη της πληροφορικής

Λαμβάνοντας υπόψη ότι τα Google Trends είναι δεδομένα που προέρχονται από μία μηχανή αναζήτησης, μπορεί κανείς να καταλάβει το λόγο που το συγκεκριμένο πεδίο παρουσιάζει μεγάλη συνάφεια και προσελκύει το ερευνητικό ενδιαφέρον. Για την ακρίβεια, οι έρευνες που σχετίζονται με τα Google Trends ξεκίνησαν αμέσως μόλις κυκλοφόρησε η υπηρεσία Trends, σε συνέχεια της έρευνας που ασχολούταν με την εξέλιξη της τεχνολογίας γύρω από τις μηχανές αναζήτησης και την ανάπτυξη υπηρεσιών αναζήτησης στο διαδίκτυο (Jun et al., 2018).

Αρχικά, ο Waller (2011) επισήμανε τις πιθανές χρήσεις των πληροφοριών που παρέχονται μέσω των Google Trends και ισχυρίστηκε ότι μια μηχανή αναζήτησης δεν

είναι μόνο μια διασύνδεση για πληροφορίες ή μια συντόμευση για πρόσβαση σε ιστοσελίδες. Υποστήριξε, επίσης, ότι δεν υπάρχει στατιστικά σημαντική διαφορά στον τύπο αναζήτησης ή τα θέματα από διαφορετικές υπό-ομάδες του πληθυσμού στο διαδίκτυο (Waller, 2011). Το 2012, οι Prakash et al. (2012) ισχυρίστηκαν ότι όταν υπάρχουν δύο ανταγωνιστικά προϊόντα που διατίθενται σε ένα δίκτυο, οι πληροφορίες που παρέχονται από τα Google Trends μπορούν να χρησιμοποιηθούν για να προβλέψουν ποιο προϊόν θα κερδίσει και θα καταλάβει μεγαλύτερο μερίδιο αγοράς (Prakash et al., 2012).

Οι Vaughan και Romero-Frias (2014) χρησιμοποίησαν τα δεδομένα που πήραν μέσω των Google Trends για να ελέγξουν εάν μπορούν να προβλέψουν την ακαδημαϊκή φήμη. Εξέτασαν πανεπιστήμια από δύο διαφορετικές χώρες, τις ΗΠΑ και την Ισπανία και κατέληξαν στο συμπέρασμα ότι υπάρχει σημαντική συσχέτιση μεταξύ του όγκου αναζήτησης για το όνομα ενός πανεπιστημίου και της ακαδημαϊκής του φήμης. Γενικά, τα πιο διάσημα πανεπιστήμια, δηλαδή αυτά που βρίσκονται πιο ψηλά στη παγκόσμια λίστα κατάταξης, προσέλκυαν περισσότερο το ενδιαφέρον των χρηστών του διαδικτύου (Vaughan & Romero-Frias, 2014). Στη συνέχεια, ο Vaughan μαζί με τον Chen (2015) επέκτειναν την έρευνα τους μελετώντας τους παράγοντες που επηρεάζουν τις πηγές δεδομένων των πληροφοριών που ζητούνται μέσω του διαδικτύου συγκρίνοντας δύο από τις μεγαλύτερες μηχανές αναζήτησης του κόσμου. Πιο συγκεκριμένα, για την έρευνα αυτή χρησιμοποίησαν τα αποτελέσματα που πήραν από αναζήτηση που πραγματοποίησαν μέσω των Google Trends με αυτά που πήραν από αναζήτηση μέσω του Baidu Index. Παρά τις διαφορές μεταξύ τους όσον αφορά την τεχνολογία, όπως για παράδειγμα οι διαφορετικές μέθοδοι επεξεργασίας της γλώσσας, ο συνδυασμός των δύο πηγών δεδομένων δεν βελτίωσε την ακρίβεια της πρόβλεψης. Ωστόσο, υπήρξε μια σημαντική διαφορά μεταξύ τους από την άποψη της διαθεσιμότητας των δεδομένων. Ο Baidu Index ήταν σε θέση να παρέχει περισσότερα δεδομένα από ότι τα Google Trends. Η ανάλυσή τους έδειξε ότι το μειονέκτημα των Google Trends οφείλεται στο γεγονός ότι ο αριθμός των χρηστών που χρησιμοποιούν το Google στην Κίνα είναι σχετικά μικρός (Vaughan and Chen, 2015). Στο ίδιο πνεύμα, υπάρχει μια μελέτη που κάνει σύγκριση μεταξύ των Google Trends και του Naver Trend. Μέσω του Naver Trend μπορεί κανείς να μελετήσει στατιστικά στοιχεία που προέρχονται από μία από τις κύριες μηχανές αναζήτησης που χρησιμοποιούνται στην Κορέα. Η συγκεκριμένη έρευνα εισαγάγει την αναλογική μέθοδο για την πρόβλεψη της ζήτησης, η οποία συγκρίνει τα δεδομένα των Google Trends και του Naver Trend και δημιουργεί αναλογίες. Σε αυτή τη διαδικασία, υπήρξε επίσης μια σημαντική συσχέτιση μεταξύ του όγκου αναζήτησης των Google

Trends και του όγκου αναζήτησης του Naver Trend για ένα πανομοιότυπο προϊόν και υποστηρίχθηκε ότι υπήρξε χρονική υστέρηση στον όγκο αναζήτησης για ορισμένα από τα προϊόντα. Η πιο σημαντική συνεισφορά αυτής της μελέτης είναι ότι παρουσιάζει τη δυνατότητα αξιοποίησης των μηχανών αναζήτησης υπό μία νέα διάσταση, της πρόβλεψης κατά αναλογία (Jun et al., 2017).

Υπάρχουν όμως έρευνες όπως αυτή του Vaughan & Yang (2013) η οποία δεν επικεντρώνεται μόνο στις θετικές πτυχές των Google Trends. Αναλυτικότερα, στη συγκεκριμένη έρευνα επέλεξαν διάφορα πανεπιστήμια και επιχειρήσεις από τις ΗΠΑ και την Κίνα και συνέλεξαν δεδομένα επισκεψιμότητας ιστού (Web traffic data) από τρεις (3) πηγές δεδομένων: ίντερνετ Alexa (Alexa Internet), Google Trends και Compete. Βρήκαν ότι υπάρχει σημαντική συσχέτιση μεταξύ των δεδομένων επισκεψιμότητας ιστού και της απόδοσης, ειδικότερα της ακαδημαϊκής ποιότητας για τα πανεπιστήμια και διάφορων οικονομικών μεταβλητών για τις επιχειρήσεις. Εξέτασαν, επίσης, τα χαρακτηριστικά των τριών παραπάνω πηγών δεδομένων και τα σύγκριναν μεταξύ τους προκειμένου να καταλήξουν στο συμπέρασμα «Ποιο είναι περισσότερο χρήσιμο;». Τελικά, αποφάνθηκαν ότι τα Google Trends υπό τις συγκεκριμένες συνθήκες δεν προέβλεψαν τόσο καλά τις ζητούμενες μεταβλητές εν συγκρίσει με τους άλλους ιστότοπους, ενώ το ίντερνετ Alexa έδωσε τα καλύτερα αποτελέσματα (Vaughan & Yang, 2013).

Κοιτάζοντας τις παραπάνω μελέτες, διαπιστώνουμε ότι το ενδιαφέρον των ερευνητών δεν επικεντρώθηκε μόνο στη χρήση δεδομένων που παρέχει η εφαρμογή Google Trends, αλλά αξιολόγησαν και σύγκριναν την απόδοση της Google ως μηχανή αναζήτησης εντοπίζοντας ταυτόχρονα βελτιώσεις που θα μπορούσαν να πραγματοποιηθούν. Ορισμένες μελέτες επικεντρώθηκαν στις μηχανές αναζήτησης και σύγκριναν τη Google με παρόμοιους ιστότοπους από άλλες χώρες όσον αφορά τα δεδομένα αναζήτησης που παρείχαν, δίνοντας με αυτό τον τρόπο την ευκαιρία να ελεγχθούν τα όρια και οι δυνατότητες των πληροφοριών που παρέχονται μέσω των Google Trends.

Το πιο σημαντικό χαρακτηριστικό που προκύπτει από τη βιβλιογραφική ανασκόπηση του πεδίου των πληροφοριακών συστημάτων και γενικότερα της επιστήμης της πληροφορικής είναι ότι οι χρήστες της εφαρμογής Google Trends ενδιαφέρονται περισσότερο για τη αναζήτηση μέσω της Google παρά για οποιαδήποτε άλλη υπηρεσία ή πηγή δεδομένων. Προτιμούν να μελετήσουν τις τεχνολογίες ως έχουν χωρίς να κάνουν χρήση των δεδομένων (Jun et al., 2018).

4.2. Ιατρική και Βιοεπιστήμες

Η μελέτη των Ginsberg et al. (2009) αποτέλεσε το έναυσμα να ξεκινήσουν πολλοί ερευνητές να ενδιαφέρονται για τις δυνατότητες αξιοποίησης των Google Trends στην επιστήμη της ιατρικής. Μέσω της μελέτης τους κατάφεραν να προβλέψουν τα τρέχοντα επίπεδα γρίπης χρησιμοποιώντας δεδομένα από σχετικά ερωτήματα που είχαν γίνει στη μηχανή αναζήτησης της Google. Όπως γίνεται αντιληπτό, η συμβολή της συγκεκριμένης έρευνας στην ιατρική επιστήμη είναι μεγάλη, καθώς το αποτέλεσμα ήταν να αναπτυχθεί ένα υπολογιστικό μοντέλο (computer model) που χρησιμοποιεί δεδομένα αναζήτησης που δεν έχουν υποστεί καμία επεξεργασία (raw search data), μετατρέποντας τα σε ένα σύστημα παρακολούθησης που προβλέπει τη δραστηριότητα των ιών της γρίπης σε πραγματικό χρόνο σε μία έως δύο εβδομάδες νωρίτερα από την έκθεση που εκδίδουν τα Κέντρα Πρόληψης και Ελέγχου Νόσων (CDC) (Ginsberg et al., 2009). Μάλιστα, οι Cha et al. (2010) επηρεασμένοι από την παραπάνω μελέτη, ξεκίνησαν την έρευνα και για άλλες εφαρμογές όπως το Twitter, το οποίο έχει πολλούς χρήστες ανά τον κόσμο και ανάλογα με το τι είναι στην επικαιρότητα κάθε φορά, δημιουργεί διάφορες τάσεις. Πιο συγκεκριμένα, ανέλυσαν την επιρροή των χρηστών του Twitter χρησιμοποιώντας τρία μέτρα, με τα οποία εξέταζαν το θέμα από διαφορετική οπτική γωνία κάθε φορά: τα retweets, τις αναφορές (mentions) & το μέτρο «indegree» που αναγνωρίζει τους χρήστες που λαμβάνουν μεγάλο ενδιαφέρον και αλληλοεπιδρούν με το κοινό τους (Cha et al., 2010).

Σε συνέχεια της μελέτης των Ginsberg et al. (2009), οι Pelat et al. (2009) αποπειράθηκαν να ερευνήσουν τη δυνατότητα παρακολούθησης άλλων τριών μολυσματικών ασθενειών εκτός της γρίπης στην Γαλλία. Διαπίστωσαν ότι για καθεμία από τις τρεις μολυσματικές ασθένειες, ένα καλά επιλεγμένο ερώτημα αρκεί για να παρέχει μια χρονολογική σειρά αναζητήσεων που συσχετίζονται σε μεγάλο βαθμό με τη συχνότητα εμφάνισης. Έδειξαν, επίσης, ότι όταν η λέξη-κλειδί που πληκτρολογούσε ο χρήστης στην αναζήτηση στο Διαδίκτυο είχε επιλεγεί σωστά, ήταν ακόμη δυνατή η παρακολούθηση της οξείας διάρροιας (acute diarrhea) και της ανεμοβλογιάς (chickenpox) σε μια μη αγγλόφωνη χώρα. Συμπέραναν, επομένως, ότι η έρευνα των Ginsberg κ.α. φαίνεται να έχει μια ευρύτερη εφαρμογή, δηλαδή πέραν της γρίπης στις Ηνωμένες Πολιτείες δείχνει ότι μπορεί να προβλέψει και άλλες μολυσματικές ασθένειες σε άλλες χώρες (Pelat et al., 2009).

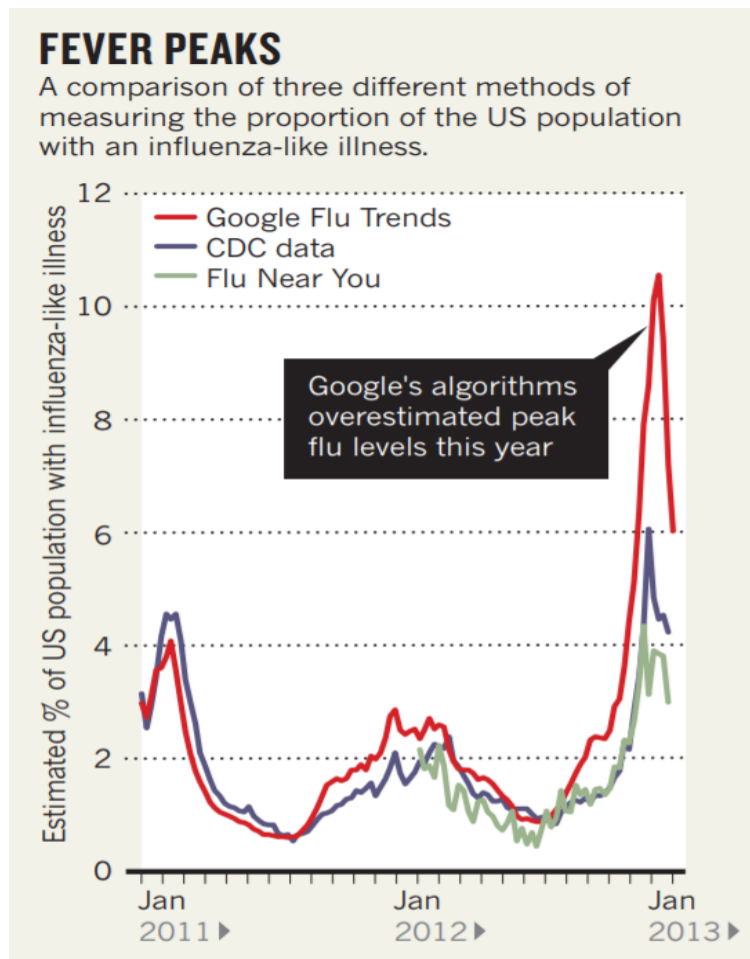
Οι Carneiro & Mylonakis (2009), επίσης, υποστήριξαν ότι η εφαρμογή των Google Trends, όσον αφορά το πεδίο της ιατρικής, αποτελεί ένα έγκαιρο και ισχυρό σύστημα παρακολούθησης, το οποίο όμως βρίσκεται σε πρώιμο στάδιο ανάπτυξης καθώς τα δεδομένα του περιέχουν μερικές ανακρίβειες. Για παράδειγμα, μπορεί κάποιος χρήστης να πληκτρολογήσει στην αναζήτηση πυρετός, κρουολόγημα και ρίγος για το ίδιο σύμπτωμα. Είναι σημαντικό λοιπόν να γνωρίζουμε κάποια χαρακτηριστικά των χρηστών αναζήτησης, όπως για παράδειγμα το επίπεδο μόρφωσης. Επιπλέον, για να είναι πιο αποτελεσματικές οι προβλέψεις μέσω των Google Trends απαιτείται μεγάλος αριθμός χρηστών να έχουν πρόσβαση στο ίντερνετ. Κατ' επέκταση, τα Google Trends παρουσιάζουν μεγαλύτερες δυνατότητες σε αναπτυγμένες χώρες (Carneiro & Mylonakis, 2009).

Οι Seifter et al. (2010) θέλησαν και αυτοί να εξερευνήσουν τις δυνατότητες που παρέχουν οι μηχανές αναζήτησης στον τομέα της υγείας χρησιμοποιώντας τη νόσο Lyme (Lyme disease). Οι λέξεις κλειδιά, ο αριθμός και η γεωγραφική τοποθεσία των αναζητήσεων μπορούν να παρέχουν δεδομένα τάσεων, έτσι όπως διατίθενται μέσω των Google Trends, και για το λόγο αυτό χρησιμοποίησαν τη συγκεκριμένη νόσο ως την πλέον κατάλληλη καθώς εμφανίζει εποχιακά και γεωγραφικά πρότυπα. Διαπίστωσαν ότι ο όγκος αναζήτησης για τον όρο “Lyme disease” αυξανόταν κατά τους εαρινούς και καλοκαιρινούς μήνες, ενώ αντίθετα ο όρος “βήχας” εμφανιζόταν κυρίως τους χειμερινούς μήνες. Επιπλέον, οι πόλεις και τα κράτη με το μεγαλύτερο όγκο αναζητήσεων για τη νόσο Lyme ταυτιζόταν σημαντικά με τις πόλεις και τα κράτη που ήταν γνωστό ότι η συγκεκριμένη νόσος είναι ενδημική. Μέσω των Google Trends, λοιπόν, κατάφεραν να προσεγγίσουν κάποιες τάσεις, τις οποίες είχαν εντοπίσει από πριν στην επιδημιολογία της νόσου Lyme, αφήνοντας να εννοηθεί ότι η δημιουργία αυτού του τύπου δεδομένων μπορεί να έχει σημαντικές θετικές επιπτώσεις στην ενίσχυση της παρακολούθησης του φάσματος των ασθενειών στο μέλλον (Seifter et al., 2010).

Υπάρχουν, επίσης, μελέτες που δείχνουν ότι τα δεδομένα που παρέχονται μέσω των Google Trends μπορούν να χρησιμοποιηθούν για την παρακολούθηση της δραστηριότητας του δάγκειου πυρετού, μίας μολυσματικής τροπικής ασθένειας που μεταδίδεται μέσω των κουνουπιών, σε διάφορες χώρες συμπεριλαμβανομένων της Ταϊλάνδης, της Βραζιλίας, της Ινδίας, της Ινδονησίας και της Σιγκαπούρης. Η έρευνα των Althouse et al. (2011) επικεντρώθηκε στη Σιγκαπούρη και τη Μπανγκόκ της Ταϊλάνδης, προσπαθώντας να προβλέψουν τη συχνότητα εμφάνισης του δάγκειου πυρετού. Τα στοιχεία των περιστατικών τα πήραν από το Υπουργείο Υγείας της

Σιγκαπούρης και το Γραφείο Επιδημιολογίας της Ταϊλάνδης και διαπίστωσαν ότι τα μοντέλα τους προβλέπουν καλά με συσχέτιση μεγαλύτερη από 0,8 (Althouse et al., 2011). Το ίδιο διαπίστωσαν στην έρευνα τους και οι Chan et al. (2011) μελετώντας τη δραστηριότητα του δάγκειου πυρετού στη Βολιβία, στη Βραζιλία, στην Ινδία, στην Ινδονησία και στη Σιγκαπούρη. Ενώ τα δεδομένα σχετικά με το δάγκειο πυρετό από επίσημες πηγές είναι συχνά διαθέσιμα μετά από σημαντική χρονική υστέρηση, τα δεδομένα αναζήτησης μέσω των Google Trends είναι διαθέσιμα σχεδόν σε πραγματικό χρόνο. Αυτά τα δεδομένα μπορούν να αποτελέσουν πολύτιμο εργαλείο στην πρόβλεψη και την παρακολούθηση της επιδημιολογίας του ιού του δάγκειου πυρετού καθώς επίσης και άλλων μολυσματικών ασθενειών σε χώρες με ανεπαρκώς αναπτυγμένα συστήματα επιτήρησης, προάγοντας με αυτό τον τρόπο τη δημόσια υγεία (Chan et al., 2011).

Παρατηρείτε ότι οι δύο παραπάνω έρευνες έρχονται σε αντιπαράθεση με την έρευνα των Carneiro & Mylonaki (2009) καθώς απέδειξαν ότι τα δεδομένα που παρέχονται μέσω της εφαρμογής Google Trends, υπό συγκεκριμένες προϋποθέσεις, μπορούν να προβλέψουν έναν ιό και σε μη αναπτυγμένες χώρες. Παρόλα αυτά, το 2013 δημοσιεύθηκε μία έρευνα που ανατρέπει όλους τους ισχυρισμούς του Ginsberg και άλλων παρόμοιων ερευνών που είχαν ανταποκριθεί θετικά στη σκοπιμότητα παρακολούθησης των τάσεων μέσω της εφαρμογής Google Flu Trends όσον αφορά τον ιατρικό τομέα. Σύμφωνα με τον Butler (2013) η Google Flu Trends υπερεκτίμησε το ποσοστό του πληθυσμού των Ηνωμένων Πολιτειών που εμφάνισε ασθένεια παρόμοια με τη γρίπη (influenza like illness - ILI) σε σύγκριση με τα δεδομένα από τα Κέντρα Ελέγχου και Πρόληψης Νοσημάτων (CDC) που βασίζονται σε αναφορές εργαστηριακής παρακολούθησης από όλες τις ΗΠΑ. Όπως φαίνεται και στο παρακάτω Σχήμα 3, η κόκκινη γραμμή, που δείχνει τις εκτιμήσεις που έκανε η Google Flu Trends, είναι περισσότερο από δύο φορές μεγαλύτερη το έτος 2013 συγκριτικά με την μπλε γραμμή που δείχνει τα πραγματικά δεδομένα έτσι όπως υπολογίστηκαν από τα Κέντρα Ελέγχου και Πρόληψης Νοσημάτων (CDC), τα οποία έχουν υπό την επίβλεψη τους 2.700 κέντρα υγείας και καταγράφουν 30 εκατομμύρια επισκέψεις ασθενών ετησίως. Παρατηρείτε ότι τις προηγούμενες χρονιές 2011 και 2012 η Google Flu Trends προέβλεψε εξαιρετικά καλά τα δεδομένα από την έκθεση του CDC σε πολλές χώρες. Αυτός ήταν ο λόγος που πολλοί ερευνητές έσπευσαν να την εκθειάσουν, το γεγονός όμως ότι το 2013 έπεσε τόσο έξω, παρά το γεγονός ότι η εφαρμογή Google Flu Trends σχεδιάστηκε για να προβλέπει τις αναφορές των CDC, δείχνει ότι οι προβλέψεις που βασίζονται στα δεδομένα αναζήτησης της Google ενδέχεται να μην είναι ακριβείς, σύμφωνα με αυτή τη μελέτη (Butler, 2013).



Σχήμα 3: Μέτρηση της γρίπης στις Ηνωμένες Πολιτείες

Πηγή: Butler, D. (2013). When Google got flu wrong. *Nature*

Ακολούθησαν και άλλες έρευνες που υποστήριξαν τα επιχειρήματα του Butler (2013), όπως αυτή των Lazer et al. (2014) που επεσήμαναν περαιτέρω ότι το πρόβλημα δεν περιορίζεται στην εφαρμογή Google Flu Trends και έθεσαν θεμελιώδη ερωτήματα σχετικά με τα είδη των φαινομένων που μπορούν να προβλέψουν τα δεδομένα αναζήτησης ή τα κοινωνικά μέσα. Υποστήριξαν, επίσης, ότι παρόλο που προηγούμενες μελέτες είχαν δείξει μερικές από τις δυνατότητες των πληροφοριών αναζήτησης, τα αποτελέσματα ανάλυσης των κοινωνικών δεδομένων (συμπεριλαμβανομένων των δεδομένων αναζήτησης ιστού) δεν είναι αρκετά αξιόπιστα για να αντικαταστήσουν τις συμβατικές μεθόδους που παραμένουν πολύτιμα εργαλεία στα χέρια των ερευνητών. Επιπλέον, ανακάλυψαν ότι δύο είναι τα προβλήματα που οδήγησαν στην εσφαλμένη πρόβλεψη του 2013: η δυναμική του αλγορίθμου και η μεγάλη βάση δεδομένων (big data hubris) με την οποία θέλουν να υποδηλώσουν την υπόθεση ότι τα μεγάλα δεδομένα (big data) τα μεταχειρίζονται οι ερευνητές ως υποκατάστατα και όχι συμπληρωματικά στην παραδοσιακή συλλογή και ανάλυση δεδομένων. Γενικά, αναγνωρίζουν τα δυνατά σημεία

της ανάλυσης και των χρήσεων των μεγάλων δεδομένων (big data) απλά εφιστούν την προσοχή όσον αφορά τη χρήση τους παράλληλα με παραδοσιακές μεθόδους (Lazer et al., 2014).

Ο τομέας της ιατρικής και της βιοεπιστήμης είναι σχετικά νέος όσον αφορά τις πληροφορίες αναζήτησης και την αξιοποίηση των δεδομένων που παρέχονται μέσω της εφαρμογής Google Trends. Οι Jun et al. (2018) θεωρούν ότι ο τομέας αυτός είναι στενά συνδεδεμένος και εμπνευσμένος από τους Ginsberg et al. (2009), καθώς μετά από έρευνα που έκαναν παρατήρησαν μεγάλη αύξηση των σχετικών μελετών μετά το 2011. Ο τομέας αυτός συνέβαλε, επίσης, σημαντικά στην αύξηση των ερευνών που σχετίζονται γενικότερα με την εφαρμογή Google Trends από το 2011 και μετά (Jun et al., 2018).

4.3. Οικονομία, Διοίκηση Επιχειρήσεων & Χρηματοοικονομικά

Στον τομέα της οικονομίας και των επιχειρήσεων, η έρευνα σχετικά με τη συμπεριφορά αναζήτησης πληροφοριών ήταν ενεργή πολύ πριν ξεκινήσει να λειτουργεί η υπηρεσία Google Trends. Η έρευνα στο συγκεκριμένο τομέα επικεντρώθηκε περισσότερο στην εφαρμογή των δεδομένων, σε αντίθεση με τον τομέα της πληροφορικής και των πληροφοριακών συστημάτων όπου η έρευνα είχε επικεντρωθεί κυρίως στις μηχανές αναζήτησης. Αυτός ήταν και ο λόγος, που η υπηρεσία Trends στον τομέα της οικονομίας γνώρισε ιδιαίτερη άνθιση από την αρχή της έναρξης της (Jun et al., 2018).

Παρακάτω αναλύονται μερικές από τις σημαντικότερες έρευνες που έγιναν όσον αφορά το πεδίο της Οικονομίας, της Διοίκηση Επιχειρήσεων & των Χρηματοοικονομικών, χωρισμένες ανά διάφορες κατηγορίες.

4.3.1. Ανεργία

Πολλές μελέτες έχουν ασχοληθεί ενεργά με τη δυνατότητα παρακολούθησης μακροοικονομικών μεγεθών μέσω της υπηρεσίας Google Trends, αρκετές εκ των οποίων επικεντρώνονται στο ποσοστό ανεργίας και στους οικονομικούς δείκτες. Οι Askitas & Zimmermann (2009) ήταν από τους πρώτους που ενδιαφέρθηκαν για τις πληροφορίες αναζήτησης που παρέχονται μέσω της Google ως μέσο γρήγορης παρακολούθησης των οικονομικών κρίσεων, καθώς είχε μόλις ξεσπάσει η μεγάλη κρίση του 2008. Πίστευαν πως η τρέχουσα οικονομική κρίση απαιτεί γρήγορες πληροφορίες προκειμένου να

προβλέψουν την οικονομική συμπεριφορά εγκαίρως, το οποίο είναι δύσκολο να συμβεί σε καιρούς διαρθρωτικών αλλαγών. Για το σκοπό αυτό, δούλεψαν και απέδειξαν ότι υπάρχει μεγάλη συσχέτιση μεταξύ των λέξεων-κλειδιών που πληκτρολογεί ένας χρήστης στην αναζήτηση και των ποσοστών ανεργίας της Γερμανίας, χρησιμοποιώντας δεδομένα δοσμένα σε μηνιαία βάση. Το καινοτόμο στοιχείο που ανέδειξε η συγκεκριμένη έρευνα είναι ότι τα δεδομένα των Google Trends μπορούν να χρησιμοποιηθούν για να προβλέψουν την οικονομική συμπεριφορά υπολογισμένα με παραδοσιακές στατιστικές μεθόδους (Askitas & Zimmermann, 2009).

Οι Choi & Varian (2009) ασχολήθηκαν, επίσης, με την ανεργία και πιο συγκεκριμένα με τις αιτήσεις παροχής επιδόματος ανεργίας (initial claims) στις Ηνωμένες Πολιτείες. Έδειξαν ότι με τα δεδομένα των Google Trends μπορεί κανείς να προβλέψει με αρκετά μεγάλη ακρίβεια την ανεργία, επτά (7) ημέρες νωρίτερα πριν από την δημοσίευση της (Choi & Varian, 2009).

Οι D'Amuri & Marcucci (2010) πρότειναν τη χρήση ενός δείκτη αναζήτησης εργασίας, του Google Index, ως τον καλύτερο δείκτη για την πρόβλεψη του ποσοστού ανεργίας των ΗΠΑ. Ο Tefft (2011) υποστήριξε, επίσης, ότι τα Google Trends θα μπορούσαν να χρησιμοποιηθούν για την πρόβλεψη πληροφοριών σχετικά με την ανεργία και έδειξε πως μία εκτεταμένη περίοδος υψηλών ποσοστών ανεργίας συσχετίζεται με υψηλά ποσοστά κατάθλιψης και άγχους. Επιπλέον, οι Fondeur & Karamé (2013) εξέτασαν και βρήκαν ότι η χρήση των δεδομένων της Google μπορεί να βελτιώσει την ακρίβεια πρόβλεψης της ανεργίας των νέων μεταξύ 15-24 ετών στην Γαλλία.

Παρόμοιες έρευνες, όσον αφορά την πρόβλεψη της ανεργίας ή των οικονομικών κρίσεων με τη χρήση των Google Trends, αποτέλεσαν σημαντικό θέμα για έρευνα και από άλλους πολλούς ερευνητές, όπως για παράδειγμα από τους Vicente et al. (2015).

4.3.2. Κατανάλωση

Οι Vosen & Schmidt (2011) σύγκριναν τον δείκτη καταναλωτικού κλίματος του Πανεπιστημίου Michigan (University of Michigan Consumer Sentiment Index – MCSI) και τον δείκτη εμπιστοσύνης των καταναλωτών (Consumer Confidence Index), ο οποίος υποδεικνύει τις τάσεις των Αμερικανών καταναλωτών, με δεδομένα που παρέχονται μέσω της υπηρεσίας Google Trends. Τα αποτελέσματα έδειξαν ότι σχεδόν σε όλα τα πειράματα πρόβλεψης, ο δείκτης της Google ξεπέρασε τους δύο άλλους δείκτες που

βασίστηκε η έρευνα. Αυτό υποδηλώνει ότι η ενσωμάτωση πληροφοριών των Google Trends σε μία μελέτη μπορεί να προσφέρει σημαντικά οφέλη σε όσους ενδιαφέρονται να προβλέψουν την ιδιωτική κατανάλωση (Vosen & Schmidt, 2011). Επιπλέον, οι Vosen & Schmidt (2012) σε μεταγενέστερη μελέτη τους, εισήγαγαν ένα νέο δείκτη που προβλέπει την ιδιωτική κατανάλωση στη Γερμανία βάσει χρονοσειρών που παρέχει η υπηρεσία Google Trends. Σε σύγκριση με τους άλλους δείκτες, ο νέος δείκτης παρείχε επίσης ακριβείς προβλέψεις στην κατανάλωση, πέραν εκείνης που έχει ήδη καταγραφεί σε άλλες μακροοικονομικές μεταβλητές (Vosen & Schmidt, 2012).

4.3.3. Χρηματιστήριο

Οι Mondria et al. (2010) μαζί με τους Da et al. (2011) ήταν οι πρώτοι που χρησιμοποίησαν δεδομένα από την αναζήτηση στο διαδίκτυο στη χρηματοοικονομική επιστήμη. Αναλυτικότερα, οι Mondria et al. (2010) ερεύνησαν τον προσδιορισμό ενός χαρτοφυλακίου χρεογράφων που βασίζεται κυρίως στην επένδυση στις εγχώριες αγορές (home bias) και του ενδιαφέροντος των επενδυτών. Χρησιμοποιώντας μία προσέγγιση καθοριστικών μεταβλητών βρήκαν εμπειρικά στοιχεία που αποδεικνύουν ότι υπάρχει σχέση αιτίας και αποτελέσματος μεταξύ του χαρτοφυλακίου χρεογράφων και της προσοχής των επενδυτών (Mondria et al., 2010).

Επιπλέον, οι Da et al. (2011) πρότειναν μία νέα μέθοδο μέτρησης του ενδιαφέροντος των επενδυτών χρησιμοποιώντας το δείκτη “Google Search Volume Index” (SVI), δείχνοντας με αυτό τον τρόπο τη χρησιμότητα των δεδομένων αναζήτησης στις χρηματοοικονομικές εφαρμογές. Ο δείκτης SVI, σε ένα δείγμα 3000 μετοχών από το 2004 έως το 2008, αποδείχθηκε πολύ ισχυρός επιτρέποντας τον γρήγορο εντοπισμό του ενδιαφέροντος των επενδυτών, ακόμη και των μικροεπενδυτών. Διαπίστωσαν ότι μία αύξηση του SVI προέβλεπε υψηλότερες τιμές των μετοχών στις επόμενες δύο (2) εβδομάδες και πιθανή αντιστροφή της τιμής εντός του έτους. Ο δείκτης SVI συνέβαλε, επίσης, σημαντικά στην απόδοση της πρώτης μέρας (first-day return) των μετοχών της εταιρίας που προσφέρθηκαν με τη διαδικασία της δημόσιας εγγραφής (Initial Public Offering – IPO), κατά την οποία ανάλογα με το ποσοστό κάλυψης της προσφοράς από τις εγγραφές προσδιορίζεται, τελικά, ο αριθμός των μετοχών που δικαιούται ο κάθε νέος μέτοχος. Τέλος, επεσήμαναν ότι ο όγκος αναζήτησης στο διαδίκτυο (internet search volume) είναι ένας αντικειμενικός τρόπος για να αποκαλυφθούν και να ποσοτικοποιηθούν τα

συμφέροντα των επενδυτών και ως εκ τούτου είναι πολύ πιθανό να υπάρχουν και άλλες εφαρμογές που θα μπορούσαν να χρησιμοποιηθούν στα χρηματοοικονομικά (Da et al., 2011).

Το 2013, οι Preis et al. εξήγησαν πως από τις τεράστιες πηγές δεδομένων που προκύπτουν ως αποτέλεσμα της αλληλεπίδρασης του ανθρώπου με το διαδίκτυο, προσφέρεται μία νέα προοπτική, η οποία εξηγεί σε πρώτη φάση τη συμπεριφορά-κίνηση των χρηματιστηριακών αγορών και δευτερευόντως συμβάλλει στην καλύτερη κατανόηση της συλλογικής ανθρώπινης συμπεριφοράς. Για το σκοπό της έρευνας, ανέλυσαν τις αλλαγές στον αριθμό των αναζητήσεων μέσω της Google, χρησιμοποιώντας 98 χρηματοοικονομικούς όρους αναζήτησης, συμπεριλαμβάνοντας μερικούς όρους που τους προτάθηκαν από την υπηρεσία “Google Sets”, ένα εργαλείο το οποίο αναγνωρίζει σημασιολογικά συγγενικές λέξεις-κλειδιά με το θέμα του εκάστοτε ενδιαφερόμενου. Από αυτή την αναζήτηση προέκυψαν μοτίβα που μπορούν να ερμηνευτούν ως «σημάδια έγκαιρης προειδοποίησης» των κινήσεων των χρηματιστηριακών αγορών. Πιο συγκεκριμένα, αναλύοντας ιστορικά δεδομένα, από τον Ιανουάριο του 2004 έως το Φεβρουάριο του 2011, εντόπισαν αύξηση του όγκου αναζήτησης για λέξεις-κλειδιά που σχετίζονται με τις χρηματοπιστωτικές αγορές πριν την πτώση του Χρηματιστηρίου (Preis et al., 2013).

Τέλος, οι Dergiades et al. (2014) εξέτασαν αν οι πληροφορίες που περιέχονται στα κοινωνικά μέσα, όπως το Twitter, το Facebook, καθώς επίσης, και τα Google Trends, μπορούν να επηρεάσουν τις χρηματοπιστωτικές αγορές. Χρησιμοποίησαν ένα πολυπαραγοντικό σύστημα (multivariate system) και εστίασαν την προσοχή τους κυρίως στις περιφερειακές χώρες της Ευρωζώνης (Ελλάδα, Ιρλανδία, Ιταλία, Πορτογαλία και Ισπανία) καθώς επίσης και σε δύο από τις βασικές χώρες της Ευρωζώνης, την Γαλλία και την Ολλανδία. Από την έρευνα τους διαπίστωσαν ότι οι συζητήσεις στα κοινωνικά μέσα και η αναζήτηση όρων, σχετικά με την ελληνική κρίση χρέους, παρείχαν σημαντικές βραχυπρόθεσμες πληροφορίες κυρίως στη διαφορά αποδόσεων των ελληνο-γερμανικών και ιρλανδο-γερμανικών κρατικών ομολόγων, ενώ επηρέασαν σε πολύ μικρότερο βαθμό τις πορτογαλικές, τις ιταλικές και τις ισπανικές αποδόσεις. Τα spreads στη Γαλλία και τις Κάτω Χώρες δεν επηρεάστηκαν. Το συμπέρασμα στο οποίο κατέληξαν οι συγγραφείς είναι ότι η αυξημένη δραστηριότητα των κοινωνικών μέσων ενημέρωσης, από την πλευρά των κακών ειδήσεων, συνεπάγεται υψηλότερο κίνδυνο, ο οποίος με τη σειρά του έχει ως αποτέλεσμα μεγαλύτερα spreads (Dergiades et al., 2014).

4.3.4. Διοίκηση Επιχειρήσεων

Σύμφωνα με τους Goel et al. (2010) ανάλογα με το τι ενδιαφέρει τους καταναλωτές να αναζητήσουν στο διαδίκτυο, μπορεί κανείς να προβλέψει τη συλλογική τους συμπεριφορά στο κοντινό μέλλον. Κατάφεραν με τη χρήση δεδομένων αναζήτησης ιστού να προβλέψουν τις εισπράξεις του box-office του πρώτου Σαββατοκύριακου για ταινίες μεγάλου μήκους, τις πωλήσεις βιντεοπαιχνιδιών για τους πρώτους μήνες κυκλοφορίας τους και την κατάταξη των τραγουδιών στη λίστα «Billboard Hot 100». Διαπίστωσαν, επίσης, ότι η παρακολούθηση της αναζήτησης των καταναλωτών στο ίντερνετ γενικά μπορεί να ενισχύσει την απόδοση των βασικών μοντέλων, ειδικά όταν υπάρχει έλλειψη σε άλλες πηγές δεδομένων ή η ακρίβεια της πρόβλεψης δεν μπορεί να βελτιωθεί σημαντικά με άλλο τρόπο (Goel et al., 2010).

Από την άλλη πλευρά, οι Choi & Varian (2012) δεν ισχυρίζονται ότι με τη χρήση των Google Trends μπορούν να προβλέψουν το μακρινό μέλλον, ισχυρίζονται όμως ότι με τη χρήση των Google Trends μπορούν να προβλέψουν το παρόν (nowcasting). Για παράδειγμα, ο όγκος αναζήτησης των πωλήσεων αυτοκινήτων τη δεύτερη εβδομάδα του Ιουνίου μπορεί να βοηθήσει στην πρόβλεψη των πωλήσεων για όλο το μήνα Ιούνιο. Έτσι, η έκθεση πωλήσεων για το συγκεκριμένο μήνα θα είναι διαθέσιμη στα στελέχη της επιχείρησης πολύ νωρίτερα από το καθιερωμένο, αφού χωρίς την χρήση των Google Trends θα κυκλοφορούσε αρκετές εβδομάδες αργότερα, μέσα στον Ιούλιο. Επίσης, στη συγκεκριμένη μελέτη δίνουν διευκρινήσεις και εξηγούν κάποιες απλές μεθόδους, με τις οποίες μπορεί κανείς να προβλέψει το παρόν χρησιμοποιώντας τα δεδομένα των Google Trends (Choi and Varian, 2012).

Οι Jun et al. (2016) υποστήριξαν ότι η χρήση της κίνησης των πληροφοριών αναζήτησης (search traffic information) που παρέχονται από την υπηρεσία Google Trends, μπορούν να συμβάλλουν στην υιοθέτηση μιας τεχνολογίας από την πλευρά των καταναλωτών ή στην αγορά ενός προϊόντος και θα μπορούσαν μάλιστα να βοηθήσουν ακόμη περισσότερο στην ανάλυση της προτιμώμενης τεχνολογίας ή των προδιαγραφών ενός προϊόντος. Για παράδειγμα, ανέλυσαν κατά πόσο το πρότυπο CAFE (CAFE standard) κίνησε το ενδιαφέρον των καταναλωτών και αν τελικά αυτό βελτίωσε στην εξοικονόμηση καυσίμων των οχημάτων. Το πρότυπο CAFE (Corporate Average Fuel Economy) είναι ένα μέτρο που υιοθετήθηκε για οικονομικότερη κατανάλωση καυσίμων και θεσπίστηκε από την κυβέρνηση των Ηνωμένων Πολιτειών σε μια προσπάθεια μετατροπής των οχημάτων, στο να γίνουν περισσότερο φιλικά προς το περιβάλλον. Τα

αποτελέσματα έδειξαν ότι το πρότυπο CAFE εκτός από το σημαντικό αντίκτυπο που είχε στη βελτίωση της οικονομίας καυσίμου, αύξησε και το μερίδιο αγοράς αντίστοιχων οχημάτων και ενίσχυσε, επίσης, την ανάπτυξη τεχνολογιών που προωθούν την εξοικονόμηση καυσίμων. Αυτό που θέλει να αναδείξει η παρούσα μελέτη είναι ότι χρησιμοποιώντας δεδομένα δημοσίως διαθέσιμα μπορεί κάποιος να αναλύσει τη συμπεριφορά των καταναλωτών και να λάβει τις απαιτούμενες αποφάσεις - στρατηγικές προς όφελος της επιχείρησης του, όπως στην προκειμένη περίπτωση, οι κατασκευαστές αυτοκινήτων (Jun et al., 2016). Επιπλέον, ο Jun σε μία μεταγενέστερη έρευνα (2017) που έκανε με τους συναδέλφους του, υποστήριξαν ότι αναλύοντας ένα δίκτυο που αποτελείται από λέξεις αναζήτησης, οι οποίες έχουν αναζητηθεί ταυτόχρονα, μπορεί να βοηθήσει τους ερευνητές να φτιάξουν ένα χάρτη σχετικά με το πώς βλέπουν οι καταναλωτές ένα προϊόν ή μια μάρκα (Jun et al., 2017).

4.3.5. Τουρισμός

Ο τουρισμός είναι ένας από τους οικονομικούς τομείς που γνώρισε ιδιαίτερο ενδιαφέρον από τους ερευνητές, όσον αφορά την πρόβλεψη του μέσω των Google Trends. Μία από τις πρώτες έρευνες που προσπάθησαν να προβλέψουν τη διαθεσιμότητα των ξενοδοχείων ήταν αυτή των Pan et al. (2012). Οι συγγραφείς προσπάθησαν να προβλέψουν τη ζήτηση για δωμάτια ξενοδοχείων και να δημιουργήσουν το καλύτερο μοντέλο πρόβλεψης με τη χρήση των Google Trends. Συμπέραναν ότι ο όγκος των δεδομένων αναζήτησης (search volume data) αποτελεί ένα πρώιμο δείκτη, που αποκαλύπτει τα ενδιαφέροντα των τουριστών και θα μπορούσε να χρησιμοποιηθεί για την πρόβλεψη της κατανάλωσης ή διαφόρων άλλων δραστηριοτήτων που θα ενδιέφεραν τους τουρίστες να κάνουν στην τοποθεσία που σκοπεύουν να ταξιδέψουν (Pan et al., 2012).

Οι Bangwayo-Skeete & Skeete (2015) ασχολήθηκαν, επίσης, με τον τομέα του τουρισμού και προσπάθησαν να προβλέψουν τις αφίξεις των τουριστών. Μέσα από την έρευνα τους εισάγουν ένα νέο δείκτη για την πρόβλεψη της ζήτησης στον τουρισμό, που κατασκευάστηκε από τα δεδομένα αναζήτησης χρονολογικών σειρών των Google Trends. Ο δείκτης αυτός βασίζεται σε μια σύνθετη αναζήτηση «ξενοδοχείων και πτήσεων» από τρεις χώρες προέλευσης σε πέντε δημοφιλείς τουριστικούς προορισμούς στην Καραϊβική. Αναλυτικότερα, ελέγχουν με ένα μοναδικό τρόπο την απόδοση της πρόβλεψης του δείκτη χρησιμοποιώντας το αυτοπαλίνδρομο μοντέλο “Autoregressive Mixed-Data Sampling” (AR-MIDAS) σε σχέση με ένα εποχικό αυτοπαλίνδρομο κινητό

μέσο “Seasonal Autoregressive Integrated Moving Average” (SARIMA) και το αυτοπαλίνδρομο υπόδειγμα (AR). Οι προβλέψεις δώδεκα μηνών αποκάλυψαν ότι το υπόδειγμα AR-MIDAS έδωσε τα καλύτερα αποτελέσματα. Αυτό υποδηλώνει ότι οι πληροφορίες που παρέχονται μέσω της υπηρεσίας Google Trends μπορούν να προσφέρουν σημαντικά οφέλη στον τομέα του τουρισμού. Ως εκ τούτου, οι υπεύθυνοι για τη χάραξη της πολιτικής και οι επιχειρηματίες, μπορούν να επωφεληθούν σε μεγάλο βαθμό από τις δυνατότητες πρόβλεψης των δεδομένων αναζήτησης (Bangwayo-Skeete & Skeete, 2015).

Αντίστοιχη έρευνα με τους Bangwayo-Skeete & Skeete έκαναν και οι Li et al. (2017), οι οποίοι αποπειράθηκαν να προβλέψουν τη ζήτηση των τουριστών για ένα προορισμό, το Πεκίνο, δημιουργώντας ένα σύνθετο δείκτη αναζήτησης μέσω ενός γενικευμένου δυναμικού μοντέλου παραγόντων (generalized dynamic factor model-GDFM). Τα ευρήματα υποδηλώνουν ότι η προτεινόμενη μέθοδος βελτιώνει την προβλεπόμενη ακρίβεια πρόβλεψης των μηνιαίων αφίξεων στο Πεκίνο σε σχέση με τα δύο άλλα μοντέλα που συγκρίνεται: ένα μοντέλο χρονοσειράς (time series models) και ένα μοντέλο με βάση ένα δείκτη που δημιουργήθηκε από την ανάλυση κάποιων βασικών παραγόντων που επηρεάζουν τον τουρισμό (Li et al., 2017).

Καθώς οι μηχανές αναζήτησης αποτελούν κορυφαίο εργαλείο για τον προγραμματισμό διακοπών, οι Dergiades et al. (2017) προσπάθησαν και αυτοί να προβλέψουν τις αφίξεις των τουριστών, αυτή τη φορά στην Κύπρο. Χρησιμοποίησαν δεδομένα σε μηνιαία βάση από το 2004 έως το 2015 και έτρεξαν δύο ελέγχους αιτιότητας (causality testing). Διαπίστωσαν ότι με σωστά επιλεγμένους τους όρους αναζήτησης (λέξεις-κλειδιά), η υπηρεσία Google Trends αποτελεί ένα χρήσιμο προγνωστικό εργαλείο για τον τομέα του τουρισμού (Dergiades et al., 2017). Επιπλέον, οι Dergiades et al. (2018) σε μεταγενέστερη μελέτη τους, ανέλυσαν πάλι μηνιαία στοιχεία τουριστών με προορισμό την Κύπρο από το 2004 έως το 2015, με τη μόνη διαφορά ότι αυτή τη φορά έτρεξαν δύο ελέγχους μη αιτιότητας (non-causality testing). Αυτό το έκαναν γιατί όταν οι τουρίστες προέρχονται από χώρες με διαφορετικές γλώσσες και διαφορετικές πλατφόρμες μηχανών αναζήτησης, η πρόβλεψη συνολικών διεθνών αφίξεων καθίσταται μια πραγματική πρόκληση για τους ερευνητές. Στην τελευταία τους μελέτη, παρατήρησαν ότι ο διορθωμένος δείκτης για διαφορετικές γλώσσες και διαφορετικές πλατφόρμες αναζήτησης, είναι καλύτερος στην πρόβλεψη της διεθνούς τουριστικής κίνησης (international visitor volumes) σε σύγκριση με το μη προσαρμοσμένο δείκτη. Η

σπουδαιότητα της συγκεκριμένης έρευνας είναι μεγάλη, καθώς είναι η πρώτη έρευνα που καταφέρνει να βρει λύση στο πρόβλημα της γλώσσας και της πλατφόρμας αναζήτησης (Dergiades et al., 2018).

Τέλος, μία ακόμη πιο πρόσφατη μελέτη είναι αυτή των Gunter et al. (2018), σύμφωνα με την οποία εξετάζετε η πιθανότητα πρόβλεψης των συνολικών τουριστικών αφίξεων σε 4 (τέσσερις) αυστριακές πόλεις, με βάση τα “LIKES” των αναρτήσεων στις σελίδες τους στο Facebook. Η ιδιαιτερότητα της συγκεκριμένης έρευνας είναι ότι συνδυάζει στο μοντέλο πρόβλεψης που χρησιμοποιεί βάσει των “LIKES”, τα Google Trends, για να ερευνήσουν αν αυτό θα βελτιώσει περισσότερο την ακρίβεια της πρόβλεψης ή όχι (Gunter et al., 2018).

4.3.6. Αγοροπωλησίες ακινήτων

Για να λάβουν αποτελεσματικές αποφάσεις, οι καταναλωτές, τα στελέχη επιχειρήσεων ή οι υπεύθυνοι χάραξης πολιτικής πρέπει να κάνουν προβλέψεις. Ωστόσο, οι περισσότερες πηγές δεδομένων, είτε προέρχονται από την κυβέρνηση είτε από ιδιωτικές επιχειρήσεις, είναι διαθέσιμες μόνο με κάποια, σημαντική συνήθως, χρονική υστέρηση. Αυτό εμποδίζει την πρόβλεψη σε πραγματικό χρόνο. Για τους λόγους αυτούς, οι Wu & Brynjolfsson, (2009) παρουσίασαν έναν πολύ ακριβή αλλά απλό τρόπο, προκειμένου να μπορούν να προβλεφθούν οι μελλοντικές επιχειρηματικές δραστηριότητες, με τη χρήση δεδομένων από μηχανές αναζήτησης. Εφαρμόζοντας τη μεθοδολογία τους για να προβλέψουν τις τάσεις των κατοικιών, διαπίστωσαν ότι ο δείκτης τους μπορούσε να προβλέψει με ακρίβεια τις μελλοντικές πωλήσεις σπιτιών καθώς και τις τιμές τους. Κατά τη διάρκεια της περιόδου δειγματοληψίας τους, κάθε ποσοστιαία αύξηση του δείκτη αναζήτησης κατοικιών προέβλεπε πωλήσεις 121.400 επιπλέον κατοικιών για το επόμενο τρίμηνο (Wu and Brynjolfsson, 2009).

Στη συνέχεια, οι Kulkarni et al. (2009) προσπάθησαν να φτιάξουν ένα δείκτη που θα μπορεί να προβλέπει τις τιμές των σπιτιών βασισμένη σε δεδομένα αναζήτησης από το διαδίκτυο. Χρησιμοποίησαν τον έλεγχο αιτιότητας του Granger και πραγματοποίησαν παλινδρομήσεις διπλής κατεύθυνσης. Η αιτιότητα του Granger στην προκειμένη περίπτωση αποδείχθηκε ένα χρήσιμο εργαλείο (Kulkarni et al., 2009).

Η πρόθεση για την αγορά ενός σπιτιού αποκαλύπτεται όταν οι δυνητικοί αγοραστές αναζητούν στο διαδίκτυο πληροφορίες για τη μελλοντική τους κατοικία. Οι Beracha & Wintoki (2013) εξέτασαν κατά πόσο η ένταση αναζήτησης στο διαδίκτυο (online search intensity) μπορεί να προβλέψει μελλοντικές διαφορές στις τιμές των σπιτιών σε μία συγκεκριμένη πόλη και διαπίστωσαν πως αυτό μπορεί να συμβεί (Beracha and Wintoki, 2013).

Τις δυνατότητες που παρέχουν τα δεδομένα αναζήτησης της υπηρεσίας Google Trends επιβεβαιώνουν, για ακόμη μία φορά, οι Alexander et al. (2014), οι οποίοι τα χρησιμοποίησαν για να προβλέψουν τους δείκτες τιμών για τα εμπορικά ακίνητα (commercial real estate). Τα αποτελέσματα τους έδειξαν ότι τα δεδομένα της Google μπορούν να χρησιμεύσουν ως ένας πρώιμος δείκτης στην αγορά των ακινήτων (Alexander et al., 2014).

4.4. Πολιτική Επιστήμη

Εκτός από τους τομείς της πληροφορικής, της ιατρικής, της διοίκησης επιχειρήσεων και των οικονομικών, έγιναν επίσης προσπάθειες να χρησιμοποιηθούν τα Google Trends στην πολιτική επιστήμη. Οι Jun et al. (2016) υποστήριξαν ότι η κοινή γνώμη των πολιτών σχετικά με την πολιτική μπορεί να αναλυθεί χρησιμοποιώντας δεδομένα της υπηρεσίας Google Trends.

Οι Mavragani et al. (2016) ισχυρίστηκαν ότι το ενδιαφέρον των πολιτών για την ρύπανση του περιβάλλοντος θα μπορούσε να ποσοτικοποιηθεί χρησιμοποιώντας πληροφορίες αναζήτησης ιστού. Πιο αναλυτικά, εξέτασαν αν η συμπεριφορά των καταναλωτών μέσω της ηλεκτρονικής αναζήτησης συσχετίζεται με τη χρήση χημικών ουσιών που ρυπαίνουν το περιβάλλον. Για την έρευνα χρησιμοποίησαν δεδομένα χρονοσειρών από την εφαρμογή Google Trends, για το χρονικό διάστημα: Ιανουάριο του 2004 έως Δεκέμβριο του 2014. Τα αποτελέσματα της παρούσας έρευνας έδειξαν ότι τα δεδομένα των Google Trends μπορούν να αποδειχθούν πολύτιμα όσον αφορά την ανάλυση της συμπεριφοράς των πολιτών καθώς επίσης θα μπορούσαν να βοηθήσουν στην αξιολόγηση και πρόβλεψη των συγκεντρώσεων των χημικών ουσιών που εξετάστηκαν (Mavragani et al., 2016).

Παρόλα αυτά, δεν ήταν τα αποτελέσματα όλων των ερευνών ελπιδοφόρα, όπως για παράδειγμα συνέβη στην έρευνα των Lui et al. (2011). Στην συγκεκριμένη έρευνα,

χρησιμοποίησαν δεδομένα των Google Trends για να προβλέψουν ποιος υποψήφιος θα κερδίσει τις εκλογές του Κογκρέσου των Ηνωμένων Πολιτειών το 2008 και το 2010. Τα αποτελέσματα όμως έδειξαν ότι δεν υπάρχει συσχέτιση μεταξύ δημοτικότητας και πιθανότητας νίκης. Με άλλα λόγια, ο υποψήφιος του οποίου το όνομα είχε αναζητηθεί περισσότερες φορές στο διαδίκτυο, δεν κέρδισε τελικά τις εκλογές του 2008. Το ίδιο συνέβη και στις εκλογές του 2010. Σύμφωνα με τους συγγραφείς, η αδυναμία πρόβλεψης οφείλεται στο γεγονός ότι μπορεί να υπάρξουν πολλές αρνητικές αναζητήσεις πληροφοριών για υποψήφιους σε τέτοιους αμφισβητούμενους τομείς, όπως αυτός της πολιτικής. Η έρευνα αυτή όμως είναι πολύ σημαντική, γιατί τα αποτελέσματα της αναδεικνύουν τα όρια της δύναμης που έχει μία πρόβλεψη που γίνεται μέσω της υπηρεσίας Google Trends (Lui et al., 2011).

5. Συμπεράσματα

Ένα μεγάλο πλεονέκτημα της υπηρεσίας Google Trends είναι ότι συλλέγει “Big Data”, επεξεργάζεται τις πληροφορίες προς διευκόλυνση της ανάλυσης και στη συνέχεια, παρέχει τις πληροφορίες αυτές δωρεάν. Ως εκ τούτου, μέσω της εφαρμογής Google Trends αναδεικνύονται με τον καλύτερο δυνατό τρόπο, οι δυνατότητες και οι περιορισμοί της χρήσης των “Big Data”, ειδικά αν αναλογιστεί κανείς ότι η μηχανή αναζήτησης της Google κατέχει το 92.37% του μεριδίου αγοράς παγκοσμίως (StatCounter, 2018). Επιπλέον, οι τάσεις, που εμφανίζει η υπηρεσία Google Trends, παρέχουν σημαντικές πληροφορίες σχετικά με την εξέλιξη αυτών των δεδομένων.

Σύμφωνα με τα αποτελέσματα των ερευνητικών εργασιών που αναλύθηκαν στην παρούσα διπλωματική εργασία, γίνεται αντιληπτό ότι υπάρχουν πολλοί τομείς στους οποίους μπορούν να χρησιμοποιηθούν τα δεδομένα αναζήτησης ιστού. Εκτός από το πεδίο της επιστήμης της πληροφορικής και των πληροφοριακών συστημάτων, το οποίο δημιούργησε μηχανές αναζήτησης, το πεδίο που χρησιμοποίησε για πρώτη φορά αυτά τα δεδομένα ήταν της ιατρικής επιστήμης. Εάν μπορεί να προβλεφθεί μία επιδημία, όπως για παράδειγμα η γρίπη, τότε η κυβέρνηση θα μπορεί να ανταποκριθεί άμεσα λαμβάνοντας μέτρα για την ενίσχυση της δημόσιας υγείας και οι φαρμακευτικές εταιρείες θα είναι σε θέση να διαχειριστούν αποτελεσματικότερα τη ζήτηση φαρμάκων που θα χρειαστούν. Στη συνέχεια, γνωρίζοντας την ζήτηση των φαρμάκων θα είναι σε θέση να προβλέψουν με μεγαλύτερη ακρίβεια τις πωλήσεις και αυτό θα τους δώσει την ευκαιρία να αυξήσουν το μερίδιο αγοράς καθώς και την κερδοφορία τους. Επιπλέον, πολλές ερευνητικές εργασίες στο πεδίο της οικονομίας και των χρηματοοικονομικών απέδειξαν πως με τη χρήση των δεδομένων μέσω της εφαρμογής Google Trends μπορεί κάποιος να προβλέψει με μεγάλη ακρίβεια τις πωλήσεις προϊόντων, κατοικιών, αυτοκινήτων, μετοχών καθώς επίσης και να προβλέψει τη συμπεριφορά των καταναλωτών. Τέλος, θα εμφάνιζε μεγάλο ενδιαφέρον η έρευνα να δραστηριοποιηθεί γύρω και από άλλους τομείς όπως η γεωγραφία, ο προγραμματισμός & η ανάπτυξη, η επικοινωνία, οι οποίοι μέχρι σήμερα έχουν προσελκύσει λιγότερο το ακαδημαϊκό ενδιαφέρον.

Τα συμπεράσματά της παρούσας εργασίας μπορεί να αποδειχθούν πολύ χρήσιμα για επιχειρήσεις που ενδιαφέρονται να προβλέψουν τις μελλοντικές τους πωλήσεις, για κυβερνητικούς οργανισμούς και υπεύθυνους για τη χάραξη πολιτικής όταν στόχος τους είναι να καταναείμουν αποτελεσματικά τους υπάρχοντες περιορισμένους πόρους και να

σχεδιάσουν βραχυπρόθεσμες ή μακροπρόθεσμες στρατηγικές προώθησης και επενδύσεων.

Βιβλιογραφία

- Alexander Dietzel, M., Braun, N. and Schäfers, W. (2014). Sentiment-based commercial real estate forecasting with Google search volume data. *Journal of Property Investment & Finance*, 32(6), pp.540-569.
- Althouse, B., Ng, Y. and Cummings, D. (2011). Prediction of Dengue Incidence Using Search Query Surveillance. *PLoS Neglected Tropical Diseases*, 5(8), p.e1258.
- Askitas, N. and Zimmermann, K. (2009). Google Econometrics and Unemployment Forecasting. *Applied Economics Quarterly*, 55(2), pp.107-120.
- Bangwayo-Skeete, P. and Skeete, R. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, 46, pp.454-464.
- Beatty, S. and Smith, S. (1987). External Search Effort: An Investigation Across Several Product Categories. *Journal of Consumer Research*, 14(1), pp.83-95.
- Beheshti-Kashi, S., Karimi, H., Thoben, K., Lütjen, M. and Teucke, M. (2014). A survey on retail sales forecasting and prediction in fashion markets. *Systems Science & Control Engineering*, 3(1), pp.154-161.
- Beracha, E., & Wintoki, M. B. (2013). Forecasting residential real estate price changes from online search activity. *Journal of Real Estate Research*, 35(3), 283-312.
- Bettman, J. and Park, C. (1980). Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision Processes: A Protocol Analysis. *Journal of Consumer Research*, 7(3), pp.234-248.
- Boone, T., Ganeshan, R., Hicks, R. and Sanders, N. (2018). Can Google Trends Improve Your Sales Forecast?. *Production and Operations Management*, 27(10), pp.1770-1774.
- Butler, D. (2013). When Google got flu wrong. *Nature*, 494(7436), pp.155-156.
- Cadet, R. L. (2018). Forecasting Consumer Price Index in Haiti using Google Trends.
- Carneiro, H. and Mylonakis, E. (2009). Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks. *Clinical Infectious Diseases*, 49(10), pp.1557-1564.

- Carrière-Swallow, Y. and Labbé, F. (2013). Nowcasting with Google Trends in an Emerging Market. *Journal of Forecasting*, 32(4), pp.289-298.
- Cha, M., Haddadi, H., Benevenuto, F. and Gummadi, K. (2010). Measuring user influence in twitter: The million follower fallacy. *Icwsn*, 10(10-17), p.e30.
- Chan, E., Sahai, V., Conrad, C. and Brownstein, J. (2011). Using Web Search Query Data to Monitor Dengue Epidemics: A New Model for Neglected Tropical Disease Surveillance. *PLoS Neglected Tropical Diseases*, 5(5), p.e1206.
- Ching-Chin, C., Ka Ieng, A., Ling-Ling, W. and Ling-Chieh, K. (2010). Designing a decision-support system for new product sales forecasting. *Expert Systems with Applications*, 37(2), pp.1654-1665.
- Choi, H. and Varian, H. (2009). Predicting initial claims for unemployment benefits. *Google Inc*, pp.1-5.
- Choi, H. and Varian, H. (2012). Predicting the Present with Google Trends. *Economic Record*, 88, pp.2-9.
- D'Amuri, F. and Marcucci, J. (2010). 'Google It!' Forecasting the US Unemployment Rate with A Google Job Search Index. *SSRN Electronic Journal*.
- Da, Z., Engelberg, J. and Gao, P. (2009). In Search of Attention. *SSRN Electronic Journal*, 66(5), pp.1461-1499.
- Dergiades, T., Mavragani, E. and Pan, B. (2017). Arrivals of tourists in Cyprus: mind the web search intensity.
- Dergiades, T., Mavragani, E. and Pan, B. (2018). Google Trends and tourists' arrivals: Emerging biases and proposed corrections. *Tourism Management*, 66, pp.108-120.
- Dergiades, T., Milas, C. and Panagiotidis, T. (2014). Tweets, Google trends, and sovereign spreads in the GIIPS. *Oxford Economic Papers*, 67(2), pp.406-432.
- Drake, M., Roulstone, D. and Thornock, J. (2011). Investor Information Demand: Evidence from Google Searches around Earnings Announcements. *SSRN Electronic Journal*.
- Elshendy, M., Colladon, A., Battistoni, E. and Gloor, P. (2017). Using four different online media sources to forecast the crude oil price. *Journal of Information Science*, 44(3), pp.408-421.

- Ettredge, M., Gerdes, J. and Karuga, G. (2005). Using web-based search data to predict macroeconomic statistics. *Communications of the ACM*, 48(11), pp.87-92.
- Fantazzini, D. and Toktamysova, Z. (2015). Forecasting German car sales using Google data and multivariate models. *International Journal of Production Economics*, 170, pp.97-135.
- Fenn, J. and Raskino, M. (2008). *Mastering the hype cycle: how to choose the right innovation at the right time*. Boston (Mass.): Harvard Business Press.
- Fissahn, J. (2001). Marktorientierte Beschaffung in der Bekleidungsindustrie. *Eine Analyse unter besonderer Be.*
- Fondeur, Y. and Karamé, F. (2013). Can Google data help predict French youth unemployment?. *Economic Modelling*, 30, pp.117-125.
- Garrison, R. and Noreen, E. (2005). Διοικητική Λογιστική. *Κλειδάριθμος*, pp.376-384.
- Geva, T., Oestreicher-Singer, G., Efron, N. and Shimshoni, Y. (2013). Do Customers Speak Their Minds? Using Forums and Search for Predicting Sales. *SSRN Electronic Journal*.
- Ginsberg, J., Mohebbi, M., Patel, R., Brammer, L., Smolinski, M. and Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457(7232), pp.1012-1014.
- Goel, S., Hofman, J., Lahaie, S., Pennock, D. and Watts, D. (2010). Predicting consumer behavior with Web search. *Proceedings of the National Academy of Sciences*, 107(41), pp.17486-17490.
- Google. (2018). *Trends Help*. [online] Available at: <https://support.google.com/trends/?hl=ko#topic=6248052> [Accessed 8 Dec. 2018].
- Google. (2018). *Επεξήγηση της αρχικής σελίδας - Βοήθεια Trends*. [online] Available at: https://support.google.com/trends/answer/6248105?hl=el&ref_topic=6248052 [Accessed 8 Dec. 2018].
- Gunter, U., Önder, I. and Gindl, S. (2018). Exploring the predictive ability of LIKES of posts on the Facebook pages of four major city DMOs in Austria. *Tourism Economics*, p.135481661879376.

- Hand, C. and Judge, G. (2012). Searching for the picture: forecasting UK cinema admissions using Google Trends data. *Applied Economics Letters*, 19(11), pp.1051-1055.
- Jun, S., Park, D. and Yeom, J. (2014). The possibility of using search traffic information to explore consumer product attitudes and forecast consumer preference. *Technological Forecasting and Social Change*, 86, pp.237-253.
- Jun, S., Sung, T. and Park, H. (2017). Forecasting by analogy using the web search traffic. *Technological Forecasting and Social Change*, 115, pp.37-51.
- Jun, S., Yoo, H. and Choi, S. (2018). Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological Forecasting and Social Change*, 130, pp.69-87.
- Jun, S., Yoo, H. and Kim, J. (2016). A study on the effects of the CAFE standard on consumers. *Energy Policy*, 91, pp.148-160.
- Khashei, M. and Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing*, 11(2), pp.2664-2675.
- Khoury, M. and Ioannidis, J. (2014). Big data meets public health. *Science*, 346(6213), pp.1054-1055.
- Kulkarni, R., Haynes, K., Stough, R. and Paelinck, J. (2009). Forecasting housing prices with Google econometrics.
- Lazer, D., Kennedy, R., King, G. and Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176), pp.1203-1205.
- Lee, W., Shih, B. and Chen, C. (2011). Retracted: A hybrid artificial intelligence sales-forecasting system in the convenience store industry. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 22(3), pp.188-196.
- Li, X., Pan, B., Law, R. and Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management*, 59, pp.57-66.
- Liu, N., Ren, S., Choi, T., Hui, C. and Ng, S. (2013). Sales Forecasting for Fashion Retailing Service Industry: A Review. *Mathematical Problems in Engineering*, 2013, pp.1-9.

- Lui, C., Metaxas, P. and Mustafaraj, E. (2011). On the predictability of the US elections through search volume activity.
- Mavragani, A., Sypsa, K., Sampri, A. and Tsagarakis, K. (2016). Quantifying the UK Online Interest in Substances of the EU Watchlist for Water Monitoring: Diclofenac, Estradiol, and the Macrolide Antibiotics. *Water*, 8(11), p.542.
- Moe, W. (2003). Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream. *Journal of Consumer Psychology*, 13(1), pp.29-39.
- Mondria, J., Wu, T. and Zhang, Y. (2010). The determinants of international investment and attention allocation: Using internet search query data. *Journal of International Economics*, 82(1), pp.85-95.
- Naccarato, A., Falorsi, S., Loriga, S. and Pierini, A. (2018). Combining official and Google Trends data to forecast the Italian youth unemployment rate. *Technological Forecasting and Social Change*, 130, pp.114-122.
- Pan, B., Chenguang Wu, D. and Song, H. (2012). Forecasting hotel room demand using search engine data. *Journal of Hospitality and Tourism Technology*, 3(3), pp.196-210.
- Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A. and Valleron, A. (2009). More Diseases Tracked by Using Google Trends. *Emerging Infectious Diseases*, 15(8), pp.1327-1328.
- Prakash, A., Beutel, A., Rosenfeld, R. and Faloutsos, C. (2012). Winner takes all: competing viruses or ideas on fair-play networks. In *Proceedings of the 21st international conference on World Wide Web*. ACM, pp.1037-1046.
- Preis, T., Moat, H. and Stanley, H. (2013). Quantifying Trading Behavior in Financial Markets Using Google Trends. *Scientific Reports*, 3(1), pp.1684.
- Raju, P., Lonial, S. and Glynn Mangold, W. (1995). Differential Effects of Subjective Knowledge, Objective Knowledge, and Usage Experience on Decision Making: An Exploratory Investigation. *Journal of Consumer Psychology*, 4(2), pp.153-180.
- Rao, A. and Sieben, W. (1992). The Effect of Prior Knowledge on Price Acceptability and the Type of Information Examined. *Journal of Consumer Research*, 19(2), p.256.
- Schaer, O., Kourentzes, N. and Fildes, R. (2018). Demand forecasting with user-generated online information. *International Journal of Forecasting*.

- Scott, L. and Varian, H. (2015). Bayesian variable selection for nowcasting economic time series. University of Chicago Press. *Economic analysis of the digital economy*. pp.119-135.
- Seifter, A., Schwarzwald, A., Geis, K. and Aucott, J. (2010). The utility of “Google Trends” for epidemiological research: Lyme disease as an example. *Geospatial health*, 4(2), p.135.
- Sekhavat, Y. and Abdollahi, P. (2016). Can Google nowcast the market trend of Iranian mobile games? In *Web Research (ICWR), Second International Conference on*, IEEE, pp.8-12.
- Shim, S., Eastlick, M., Lotz, S. and Warrington, P. (2001). An online prepurchase intentions model. *Journal of Retailing*, 77(3), pp.397-416.
- Shukla, A. and Sharma, S. (2018). A System Dynamic Approach to Analyse the Impact of Electronic Word of Mouth on Box Office Revenue. *SAMVAD*, 14, pp.19-25.
- StatCounter Global Stats. (2018). *Search Engine Market Share Worldwide / StatCounter Global Stats*. [online] Available at: <http://gs.statcounter.com/search-engine-market-share#quarterly-200901-201702> [Accessed 8 Dec. 2018].
- Sullivan, D. (2016). Google now handles at least 2 trillion searches per year. *Search Engine Land*.
- Tefft, N. (2011). Insights on unemployment, unemployment insurance, and mental health. *Journal of Health Economics*, 30(2), pp.258-264.
- Tetlock, P. (2015). The Role of Media in Finance. *Handbook of media Economics*, Vol. 1, pp.701-721.
- To, P., Liao, C. and Lin, T. (2007). Shopping motivations on Internet: A study based on utilitarian and hedonic value. *Technovation*, 27(12), pp.774-787.
- Vaughan, L. and Chen, Y. (2015). Data mining from web search queries: A comparison of google trends and baidu index. *Journal of the Association for Information Science and Technology*, 66(1), pp.13-22.
- Vaughan, L. and Romero-Frías, E. (2014). Web search volume as a predictor of academic fame: An exploration of Google trends. *Journal of the Association for Information Science and Technology*, 65(4), pp.707-720.

- Vaughan, L. and Yang, R. (2013). Web traffic and organization performance measures: Relationships and data sources examined. *Journal of Informetrics*, 7(3), pp.699-711.
- Vicente, M., López-Menéndez, A. and Pérez, R. (2015). Forecasting unemployment with internet search data: Does it help to improve predictions when job destruction is skyrocketing?. *Technological Forecasting and Social Change*, 92, pp.132-139.
- Vosen, S. and Schmidt, T. (2011). Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of Forecasting*, 30(6), pp.565-578.
- Vosen, S. and Schmidt, T. (2012). A monthly consumption indicator for Germany based on Internet search query data. *Applied Economics Letters*, 19(7), pp.683-687.
- Waller, V. (2011). Not just information: Who searches for what on the search engine Google?. *Journal of the American Society for Information Science and Technology*, 62(4), pp.761-775.
- Wu, L. and Brynjolfsson, E. (2009). The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales. *SSRN Electronic Journal*.
- Youn, S. and Cho, H. (2016). Nowcast of TV Market using Google Trend Data. *Journal of Electrical Engineering and Technology*, 11(1), pp.227-233.
- Zhao, K. and Wang, C. (2017). Sales Forecast in E-commerce using Convolutional Neural Network.
- Zimmer, J., Henry, R. and Butler, B. (2007). Determinants of the Use of Relational and Nonrelational Information Sources. *Journal of Management Information Systems*, 24(3), pp.297-331.