

פרויקט טעימות מרחבי העולם

מטבחים מהעולם כמשקפים תרבויות
ואוכלוסיות

שם הקבוצה

פרויקט טעימות מרחבי העולם – מטבחים מהעולם כמשקפים תרבויות ואוכלוסיות

פרטי הקבוצה

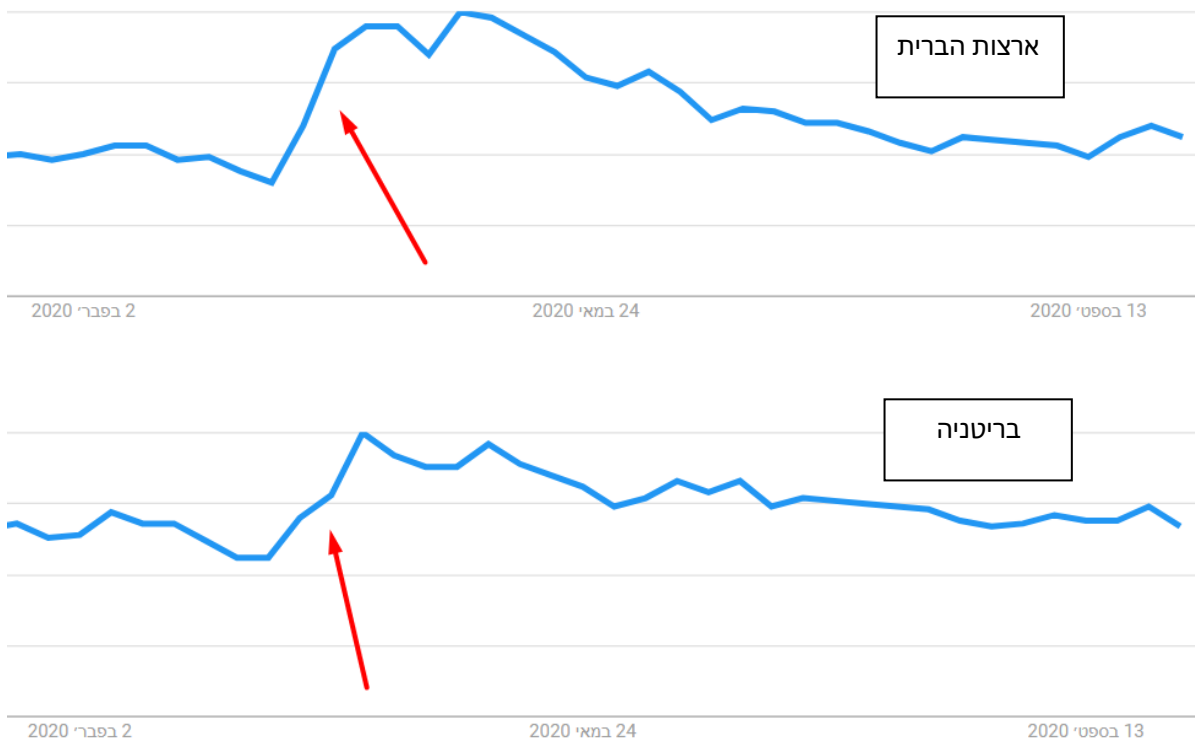
מספר הקבוצה 27

יפעת פרלמוטר, בעלת שם משתמש Yifatpe וכתובת אימייל Yifat.perlmutter@mail.huji.ac.il

אוראל בן ישראל, בעלת שם משתמש Orel.benisrael וכתובת אימייל Orel.benisrael@mail.huji.ac.il

רקע לעבודת המחקר

אוכל מהווה רכיב משמעותי ומרכזי בחייו של כל אדם באשר הוא; חשיבותו של האוכל חוצה מגזרים, חברות, אזורים דמוגרפים ותקופות לאורך ההיסטוריה האנושית. מלבד היותו רכיב הכרחי בהישרדות, מושפע האוכל גם מתהליכים חברתיים ואישיים, מתרבויות, תקופות ומצבים שונים. ניתן לראות זאת כשמתכלים ב-Google Trends אשר משקף את כמות החיפושים שנעשו עבור מילה מסוימת לאורך זמן, ורואים עלייה משמעותית בארה"ב ובבריטניה בכמות חיפושי המילה "Cake" בחודש מרץ 2020. אם נניח שמרבית חיפושים אלה נעשו לצורך חיפוש מתכון והכנת עוגה, יוצא מכך כי בתקופה בה הקורונה התפרצה בעולם חלה עלייה בהתעניינות בעוגות בארה"ב ובבריטניה. כמובן שאי אפשר להסיק סיבתיות ללא מחקר מעמיק יותר, אך אפשר לומר כי בהתאם לגרף זה, הטענה כי כשאנשים נאלצו להסתגר בבתיהם תחת חשש, אי וודאות לגבי המתרחש והפרת השגרה שלהם הם התעניינו ואולי אף הכינו ואכלו יותר עוגות תהיה ניחוש הגיוני.



לאור זמינותם של מצרכי היסוד והגיוון הרב של המאכלים, רוב בני האדם מקלים ראש בכל הנוגע למרכזיות האוכל ותופשים אותו כמובן מאליו וכדבר של מה בכך. אולם, ניתוח מעמיק ומושכל של נתונים אודותיו עשוי להעלות תובנות בעלות ערך רב. דוגמא לסוגי התובנות אשר ניתן לקבל נובעות מקורלציות וממתאמים שניתן למצוא ממאגרי מידע שונים הקשורים למאכלים השונים שניתן למצוא במטבחים האתניים השונים ברחבי העולם, לבין מאפיינים של אותה דמוגרפיה: מיקום גיאוגרפי, רמת בריאות ואושר למשל.

נתונים טכניים

במסגרת עבודת המחקר התנסינו בעבודה עם סביבות עבודה שונות: החלק הראשון של העבודה התבצעה ב-PyCharm כאשר התבססנו על שפת Python ועל מגוון ספריות כדוגמת Beautiful-Soup, Pandas, Numpy ועוד. לאחר מכן ייצאנו את הנתונים ל-Excel, השתמשנו במשוואותיו ובכלים שבו (כדוגמת Pivot-Table) ויצרנו סרטונים והצגות ויזואליות של הנתונים. בחלק השני של העבודה סביבת העבודה הייתה Jupyter Notebook כאשר התבססנו על שפת Python ועל הספריות שצוינו קודם לכן ועל ספריות matplotlib, wordcloud בעזרתן יצרנו סרטונים ותצוגות ויזואליות של הנתונים.

בהשקפתנו, זו הייתה הדרך המיטבית להרחיב את יכולת העבודה שלנו בכתיבת קוד, בסביבות עבודה שונות ואת יכולות העבודה עם כלים שונים לניתוח נתונים שאינם מבוססי קוד. בפורום הנדרש העלנו סרטון הרצה של הקודים של שלב א' של הפרויקט שלנו.

תיאור המחקר

שלב א'

בשלב הראשון רצינו לבסס את ההנחה שאוכל משתנה בין מקומות ושיש לו הקשר תרבותי וסביבתי. לשם כך רצינו לבחון האם קיימים הבדלים במאפיינים של אותו סוג מזון, בין מדינות שונות. התמקדנו באתר מתכונים בשם AllRecipes אשר לו שתי גרסאות שונות: אחת של ארה"ב והשנייה של בריטניה ואירלנד. יצאנו מתוך נקודת הנחה שאתר גדול ומרכזי של מתכונים מעיד על סגנון המאכלים במדינה וגם מושפע מהמתכונים המקובלים במדינה. באתר זה קיימים עשרות אלפי מתכונים וכדי שנוכל להשוות בין שתי מדינות בחרנו להתמקד בעוגות שוקולד, שהן עוגות מוכרות וביססיות בכל העולם המערבי, מבוססות על אותו קונספט ולהן קיימים מאות מתכונים שונים בכל אחד משני האתרים.

בנינו קראולר עבור כל אחד מהאתרים (באתר האמריקאי עבור 493 מתכונים, ובאתר הבריטי עבור 409 מתכונים), אשר שולף עבור כל מתכון את מספר המנות ואת כמויות הסוכר, החמאה והביצים. זאת, משום שסוכר, חמאה וביצים הם רכיבי בסיס בכלל ובפרט בעוגות. היו הבדלים רבים בין האתרים ולכן נדרשנו לבנות קראולר ספציפי עבור כל אחד מהם, לסנן מתכונים לא קשורים, לפתור בעיות הנובעות משינויים בתוך האתר בין מתכונים שונים, לוודא שלא מתבצע חילוק באפס ושהכמויות נמדדות באופן דומה. השתדלנו לדייק את הקוד כמה שיותר כדי שההשוואה תהיה נכונה. למשל, להמיר את כל כלי המדידה (כוס, כפית תה, כפית שולחן ועוד) מהמתכונים באתר של ארה"ב ליחס של כוסות, ומהמתכונים באתר של בריטניה ליחס של גרמים. לאחר מכן המרנו את הכוסות לגרמים גם כן כדי ליצור השוואה הולמת.

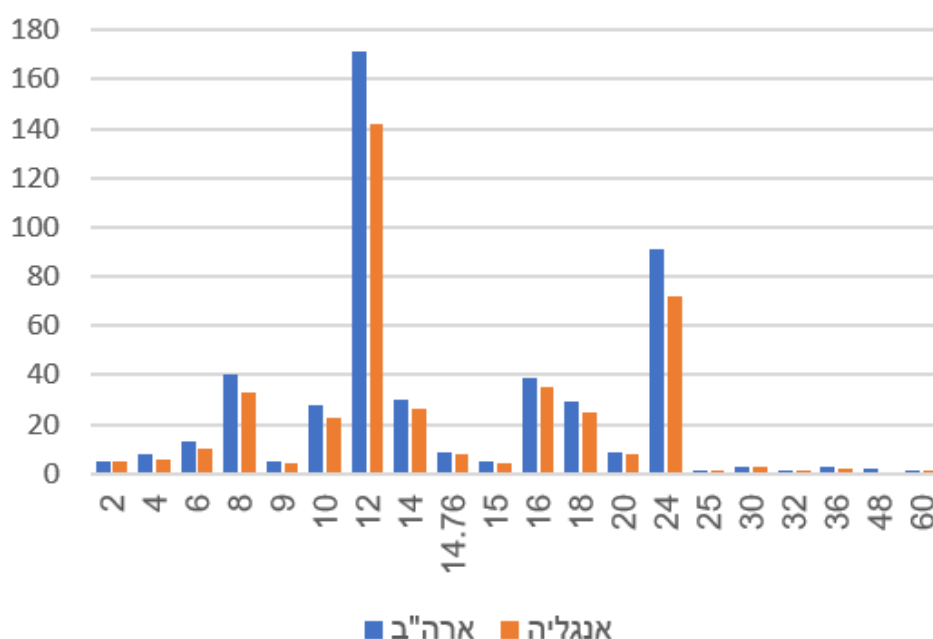
ב-Milestone ניסינו לעשות השוואה בין כמויות הסוכר בלבד, אך היא הייתה הרבה פחות מדויקת ולא לקחה בחשבון את מספר המנות המוגדר עבור כל מתכון. בפרויקט הנוכחי בנינו מחדש את הקראולרים כדי לשלוט על הלינקים ועל אופן הסינון של המידע שאספנו, דייקנו את הסינונים, הוספנו את הנתון של

מספר מנות לכל מתכון שאפשר לנו להשוות את המידע בין האתרים בצורה נכונה ולא מוטה, והוספנו את נתוני כמויות החמאה והביצים כדי להרחיב את ההשוואה ואת הבנת הדומה או השונה בין המתכונים בשתי המדינות.

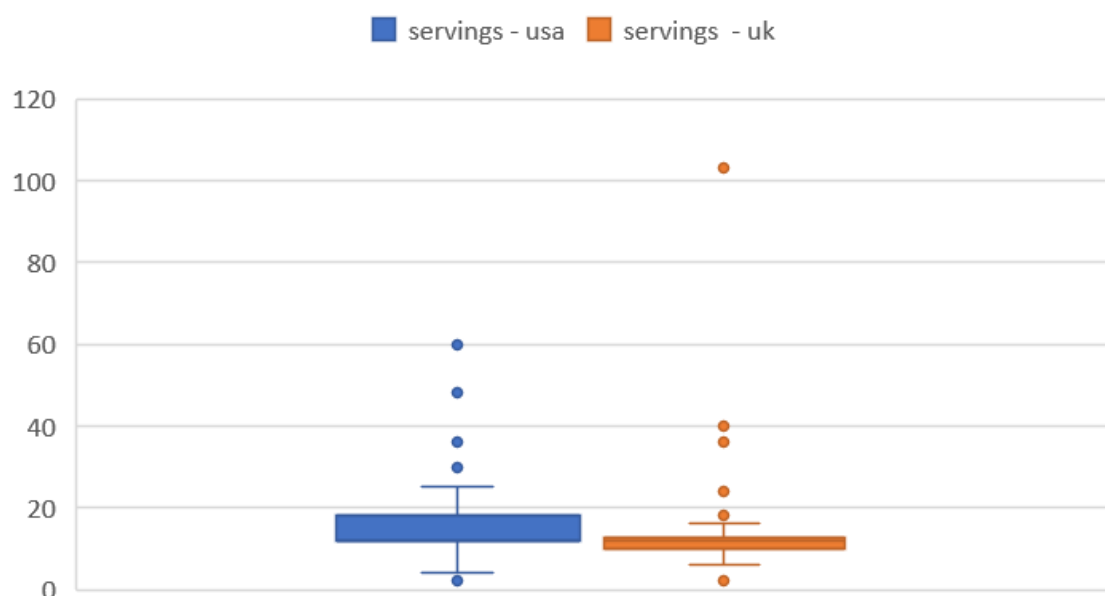
הממצאים

ראשית, הסתכלנו על התפלגות מספר המנות בכל מתכון של אנגליה וארה"ב. במבט על גרף עמודות (למטה) נראה כי ההתפלגויות דומות. לאחר מכן יצרנו גרף מסוג boxplot שבו מתחדדים ההבדלים בין ההתפלגויות וניתן לראות כי בארה"ב ממוצע המנות למתכון גבוה יותר וגם השונות גדולה יותר. נתון זה יכול להצביע על הבדלים בגדלי העוגות שנאפות. עוגה גדולה יותר עשויה לגרום לאנשים לאכול ממנה יותר, מכיוון שמהניסיון שלנו החלוקה לכמות מנות היא ספורדית כשמדובר בעוגות וחלוקת העוגה מתבצעת שונה בין אנשים. מצד שני, נתון זה יכול להצביע על כך שאותו גודל עוגה נחשב כמיועד למספר שונה של אנשים.

התפלגות מתכונים לפי מספר מנות

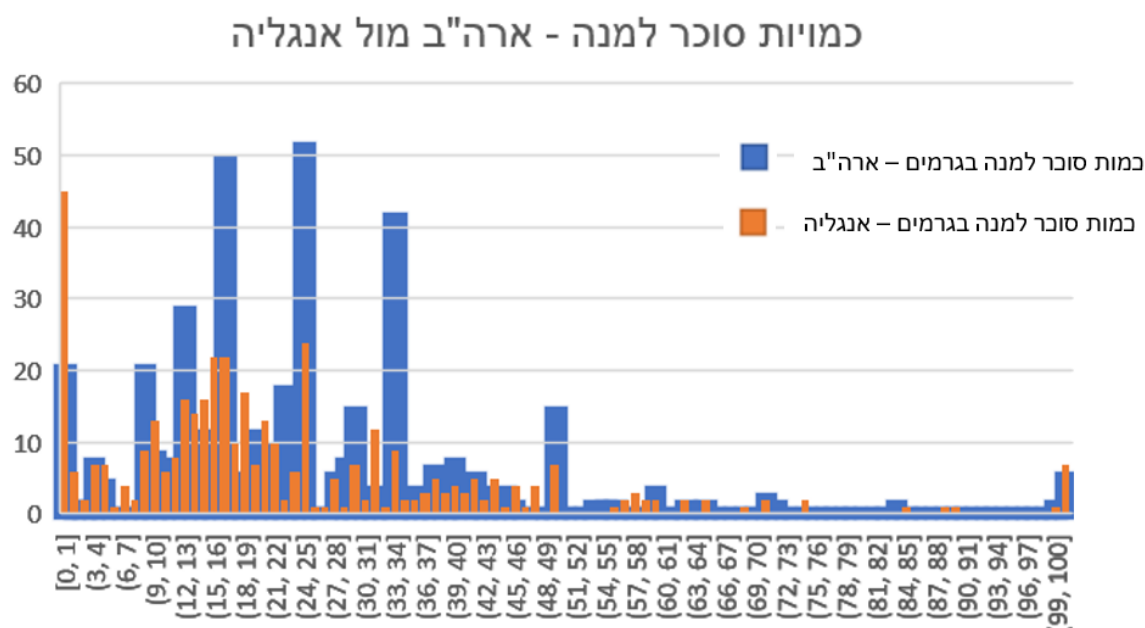


התפלגות מספר המנות - אנגליה לעומת ארה"ב

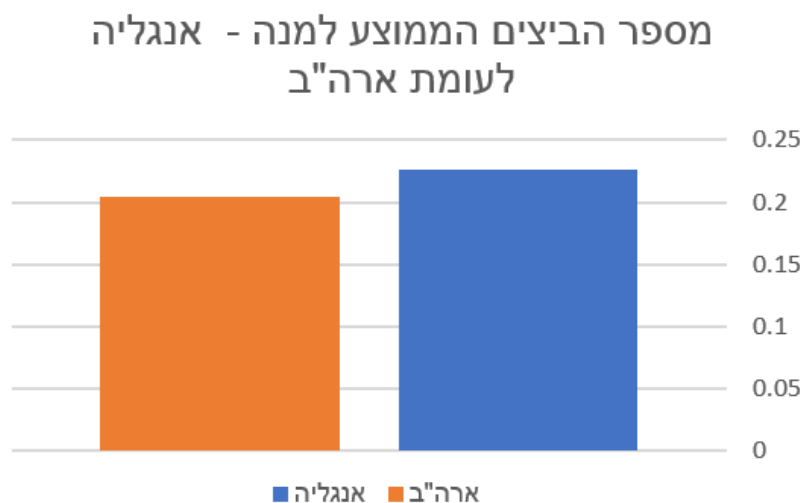


כאמור, קיבלנו מהרצת הקראולרים רשימות של כמויות חמאה, ביצים וסוכר עבור מתכוני עוגות השוקולד באתר האמריקאי ובאתר הבריטי. באמצעות Python יצרנו רשימות שמנרמלות את התוצאות לפי מספר המנות במתכון כדי שלא תיווצר הטיה של כמות רכיב מסוים בשל גודל העוגה וביצענו בדיקות וטיוב כדי לוודא שהנתונים הגיוניים, מתייחסים גם למקרי קיצון ומדויקים. ייצאנו ל-Data_Frame באקסל באמצעות pandas. לאחר מכן יצרנו באמצעות אקסל גרפים שימחישו את הדומה והשונה בין המתכונים במדינות אלה בצורה המיטבית וחישבנו ממוצעים.

גילינו כי ממוצע הסוכר למנה באתר הבריטי הוא 27.9 גרם, ובאתר האמריקאי הממוצע נמוך יותר – 25.4. ניתן לראות כי ההתפלגויות עצמן יחסית דומות (בארה"ב התצפיות ממורכזות יותר).

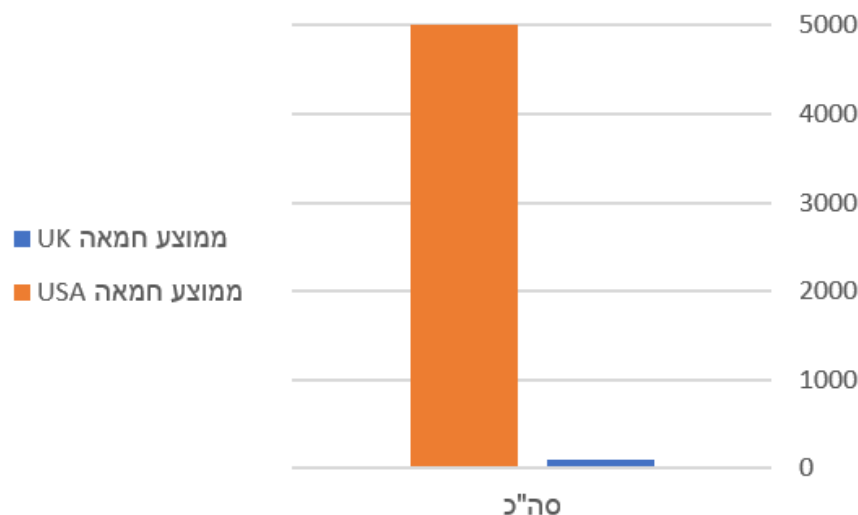


גילינו כי ממוצע כמות הביצים באתר הבריטי למנה הוא 0.225 ביצה, ובאתר האמריקאי גם כאן הממוצע נמוך יותר – 0.197.



בנוסף, התקבל ממצא מפתיע עבור ההבדלים בכמויות החמאה עבור מנת עוגה : בעוד באתר הבריטי הממוצע היה 9 גרם (ו – 103.4 גרם לעוגה), באתר האמריקאי הממוצע היה 564.3 גרם (ו – 5644 גרם לעוגה). הנתונים כוללים את הרוטב לעוגה. הבדל זה הינו קיצוני מאוד והנחנו שמדובר בטעות. בדקנו את הנתונים בכמה דרכים, וידאנו שכמות המנות לעוגה היא נכונה ושהמרת הכמויות היא נכונה, ולבסוף אכן נמצאו מספר תצפיות חריגות מאוד, שגרמו להבדל הממוצעים להראות כך :

כמות החמאה הממוצעת למנה - אנגליה לעומת ארה"ב



לאחר מחיקת התצפיות החריגות וחישוב הממוצע החדש, קיבלנו שממוצע הסוכר החדש באתר האמריקאי למנה הוא 14.1, ממוצע הגיוני הרבה יותר אך עדיין הבדל גדול מהממוצע באתר האנגלי והצעות לשיפור יהיו לדייק את הקוד במחקרים הבאים ולהגדיל את מספר התצפיות, וגם לחקור עוד את האפשרות שאכן קיים הבדל מובהק. תוצאות אלה מבססות את הטענה כי ייתכן וקיימים הבדלים בין מאכלים מאותו הסוג במדינות שונות. הבנה מעמיקה יותר של מקורן או תוצאותיהן (למשל, השמנת יתר או עודף סוכר בדם) תאפשר מלבד הבנה מעמיקה יותר של בני האדם, חוש הטעם שלהם ומקורות התזונה שלהם לפתח גם כיוונים ורעיונות לשיפור או שימור הרגלים הנפוצים בחברה לגבי תזונה. כך למשל, ייתכן כי שינוי מכוון של הכמויות המופיעות במתכונים תגרום לצריכה מופחתת של חומרים שאינם בריאים לגוף.

שלב ב'

בשלב השני של העבודה רצינו להבין לעומק הבדלים בין מאכלים ומטבחים שונים בין מדינות שונות. כפי שראינו שקיימים הבדלים ברכיבי אותם מאכלים בין מדינות שונות, קל וחומר שהגיוני להניח שקיימים הבדלים משמעותיים בסוג המאכלים במדינות השונות וברכיביהם. לכן, התחלנו בניתוח מאגר נתונים שלקחנו מ-Kaggle, מדובר בקובץ JSON של 12.4 MB. כפי שנאמר ניתוח זה התבצע בסביבת העבודה של Jupyter Notebook. במחקר בחנו 39,774 מתכונים של 20 מדינות, מדובר בכמות מידע גדולה ולכן האתגר המרכזי היה לנתח את הנתונים מפרספקטיבות שונות.

כשלב ראשוני ביצענו Reality Check לקובץ JSON אשר ניתחנו. במסגרת תהליך זה בדקנו את תכולת התיקיה ממנה שלפנו את הנתונים ; לאחר ביצוע המרה של אופן אחסון המידע מקובץ ה-JSON לטבלה ב-Jupyter Notebook בדקנו כיצד נראית דגימה קטנה ממאגר הנתונים. המאגר שיצרנו מכיל 3 עמודות ו-39774 שורות. כל שורה מכילה מידע עבור מתכון, המידע הנתון הוא מספר זיהוי למתכון, שיוכו למטבח, ורשימת רכיבים ממנו הוא מורכב.

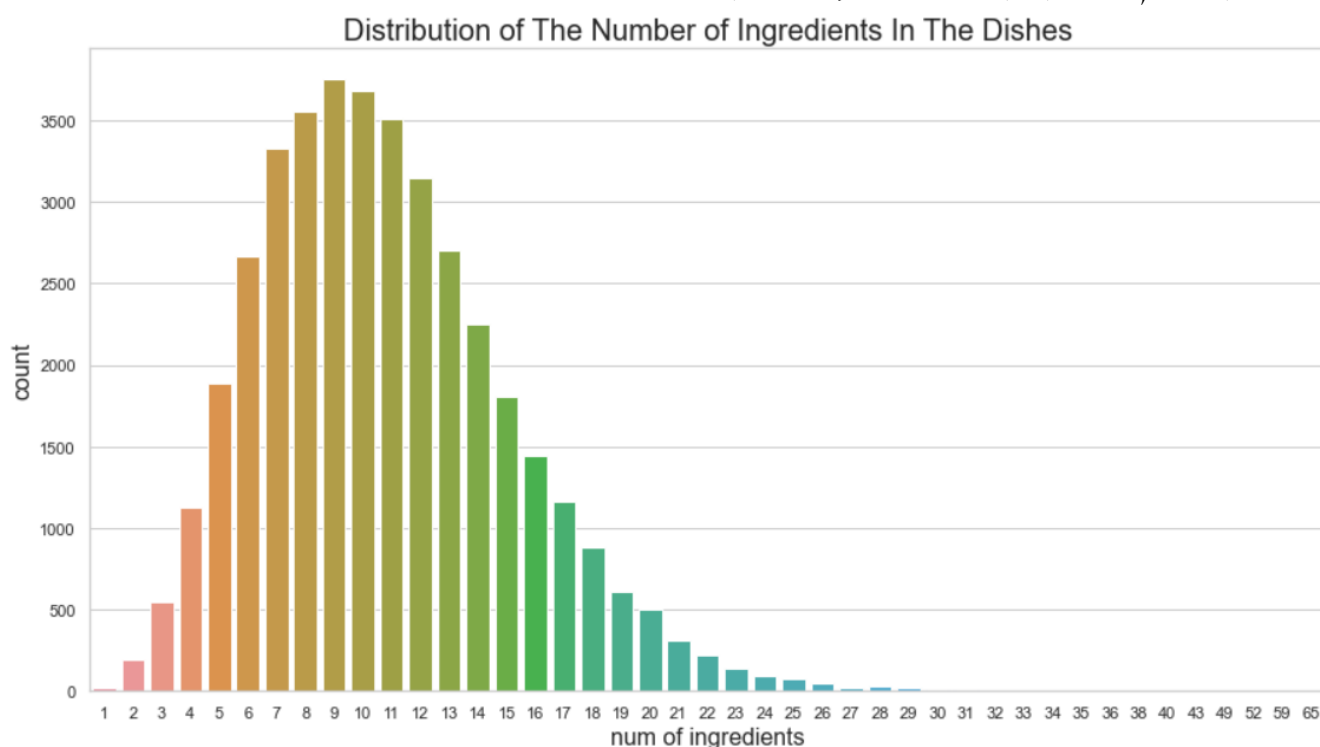
ingredients	cuisine	id	
...romaine lettuce, black olives, grape tomatoes]	greek	10259	0
...plain flour, ground pepper, salt, tomatoes, g]	southern_us	25693	1
...eggs, pepper, salt, mayonaise, cooking oil, g]	filipino	20130	2
[water, vegetable oil, wheat, salt]	indian	22213	3
...black pepper, shallots, cornflour, cayenne pe]	indian	13162	4

על מנת שנוכל לבדוק ולנתח את הנתונים ביתר קלות הוספנו לטבלת הנתונים שלנו עמודה חדשה המכילה את מספר הרכיבים במנה. הוספת עמודה זו הקלה משמעותית על המשך העבודה.

כחלק מהבדיקה של הנתונים בחנו את הכמות המקסימלית, המינימלית והממוצעת של רכיבים במנה, גילינו כי המנה בעלת הרכיבים הרבים ביותר הינה המנה האיטלקית אשר מכילה 65 רכיבים. בחנו את המילים ברשימה וראינו כי אכן כל המילים בה נכונות ותקינות, על כן לא מדובר בתצפית שגויה. לאחר מכן גילינו כי הכמות המינימלית של רכיבים במנה במאגר המידע הינה רכיב אחד. מנתון זה הסקנו כי לא קיים במאגר חוסרים של רכיבים (מנות ללא רכיבים). גילינו כי קיימות מספר מנות המכילות רכיב אחד, בחנו את הרכיבים שלהן וראינו כי אכן כולם תקינים.

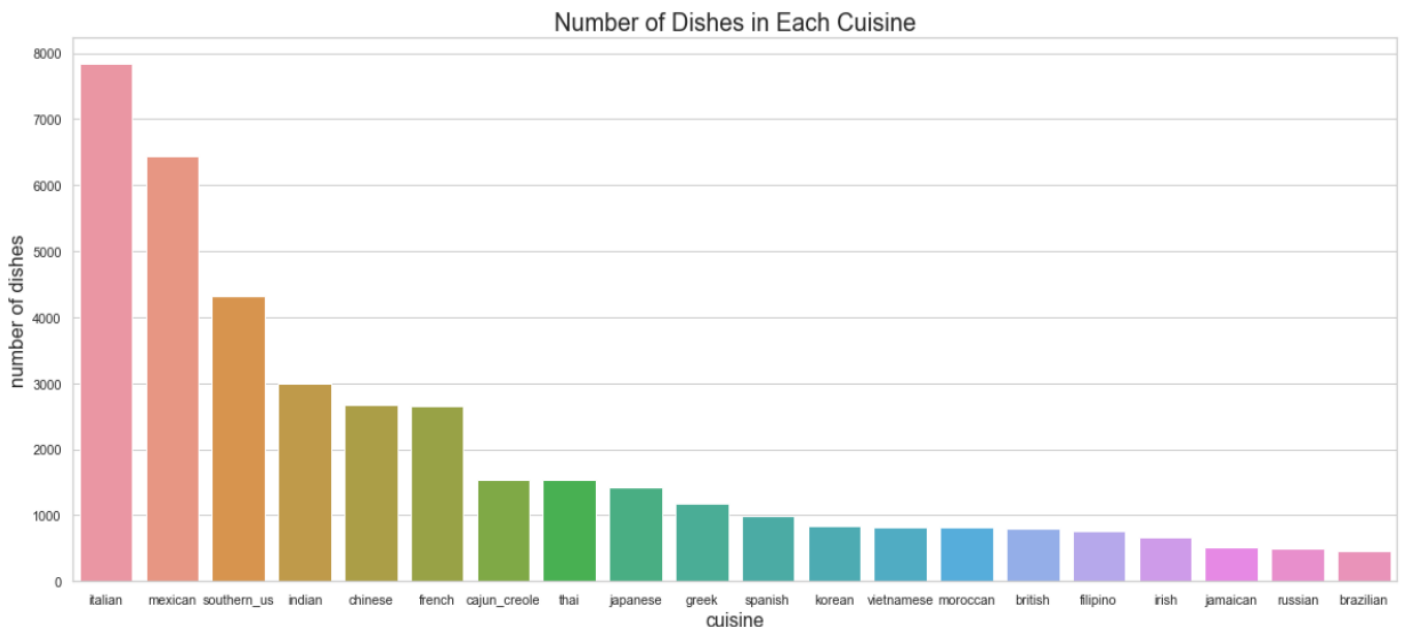
לאחר ביצוע בחינה של התצפיות הקיצוניות בדקנו את הכמות הממוצעת של רכיבים במנה בכל המטבחים, התוצאה שקיבלנו הינה כ-10.76 רכיבים במנה.

לאחר מכן, בדקנו את התפלגות כמות הרכיבים במנה כדי לאמת את התוצאות שקיבלנו בבדיקה הקודמת, במסגרת הבדיקה נוכחנו לראות כי אכן קיבלנו עקומת פעמון, כאשר כל המאכלים מכילים כמות חיובית של רכיבים, ושמספר הרכיבים למנה עם השכיחות הכי גבוהה הוא 9 :



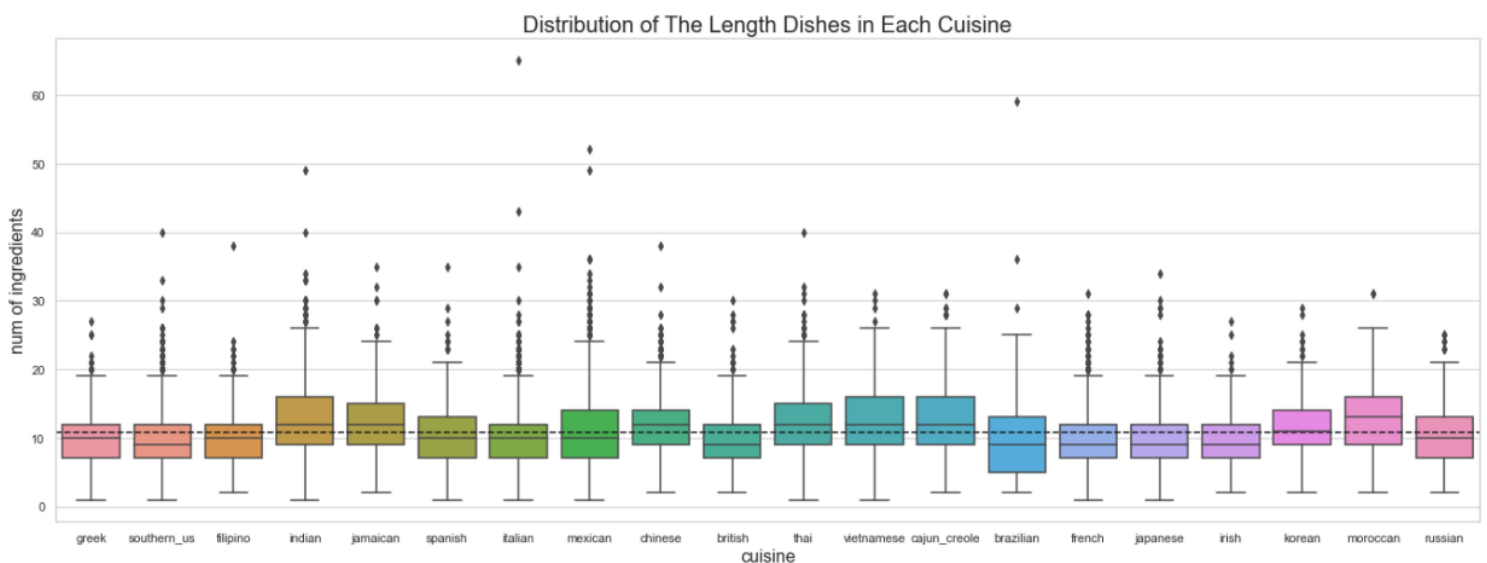
לאחר ביצוע ה-Reality Check עברנו לבדוק את הנתונים עצמם מנקודות מבט שונות.

תחילה בדקנו את כמות המתכונים של כל מטבח הקיימים במאגר ; כפי שניתן לראות רוב המתכונים הם מהמטבח האיטלקי, אחריו המקסיקני בעוד שלמטבח הברזילאי יש הכי מעט מתכונים במאגר. להלן כמות המתכונים המופיעים במאגר בחלוקה למטבחים :



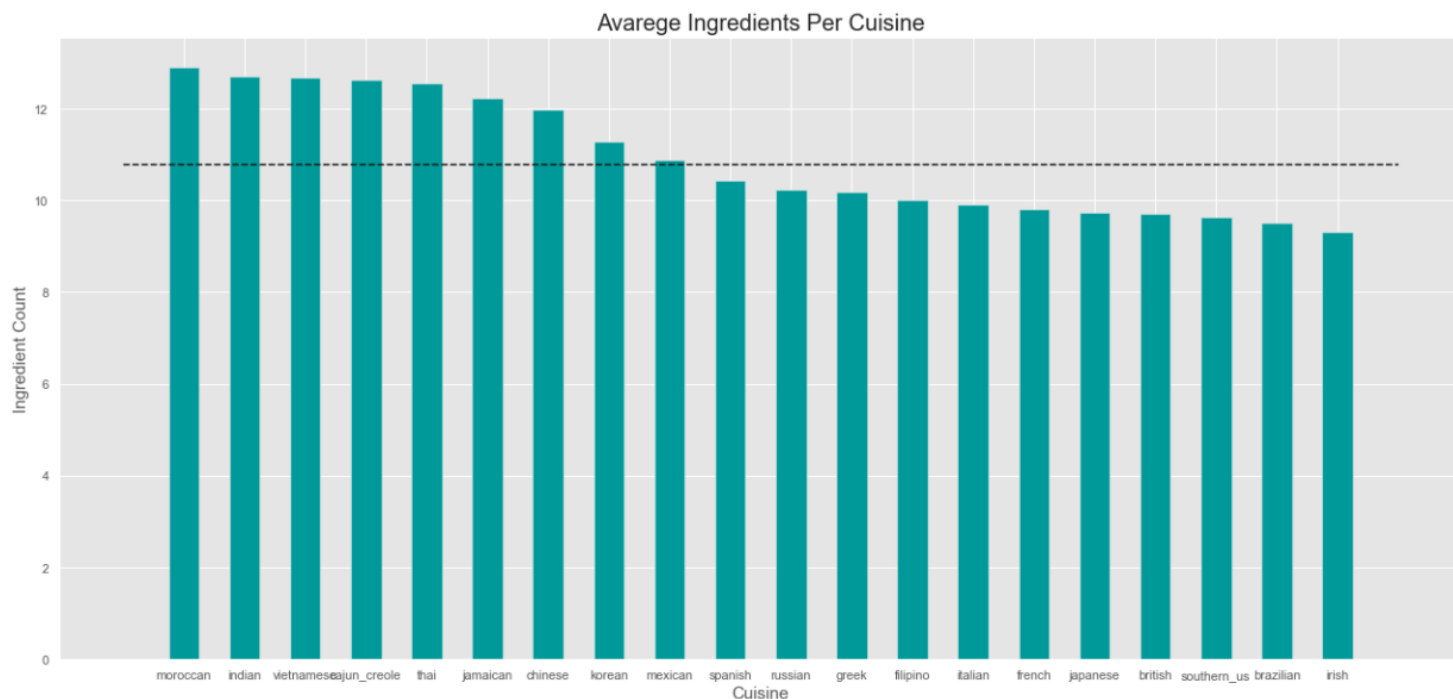
מכאן ניתן להבין כי יש לנו כמות נתונים על המתכונים במטבחים השונים שאיננה מאוזנת, דבר אשר עלול להשפיע על ניתוחים ועל שרטוטים ולגרום לנו להסיק מסקנות לא נכונות, על כן היינו צריכות לשים לב אילו נתונים מושפעים מכמות המנות השונה בין המטבחים.

לאחר מכן בחנו את התפלגות כמות הרכיבים במנות שיש בכל מטבח, לשם נוחות הוספנו ב-Box plots ישר מאוזן המייצג את ממוצע כמות הרכיבים בכל מנה, אשר נפוצים ברוב המטבחים :



מגרף זה ניתן לראות כי בכל המטבחים קיימות תצפיות חריגות, אכן התצפית הכי חריגה היא של המטבח האיטלקי של המנה בעלת 65 הרכיבים.

לאחר מכן, בחנו את הממוצעים של כמויות הרכיבים שיש במתכונים בכל אחד מהמטבחים השונים, נזכור כי ממוצע כמות הרכיבים במנה של כל המטבחים הינו כ-10.7677 רכיבים. להלן גרף המתאר את ממוצע הרכיבים שיש במנה בכל מטבח, כאשר הקו האופקי מתאר את הממוצע הכולל של מספר הרכיבים במנה בכל המטבחים:

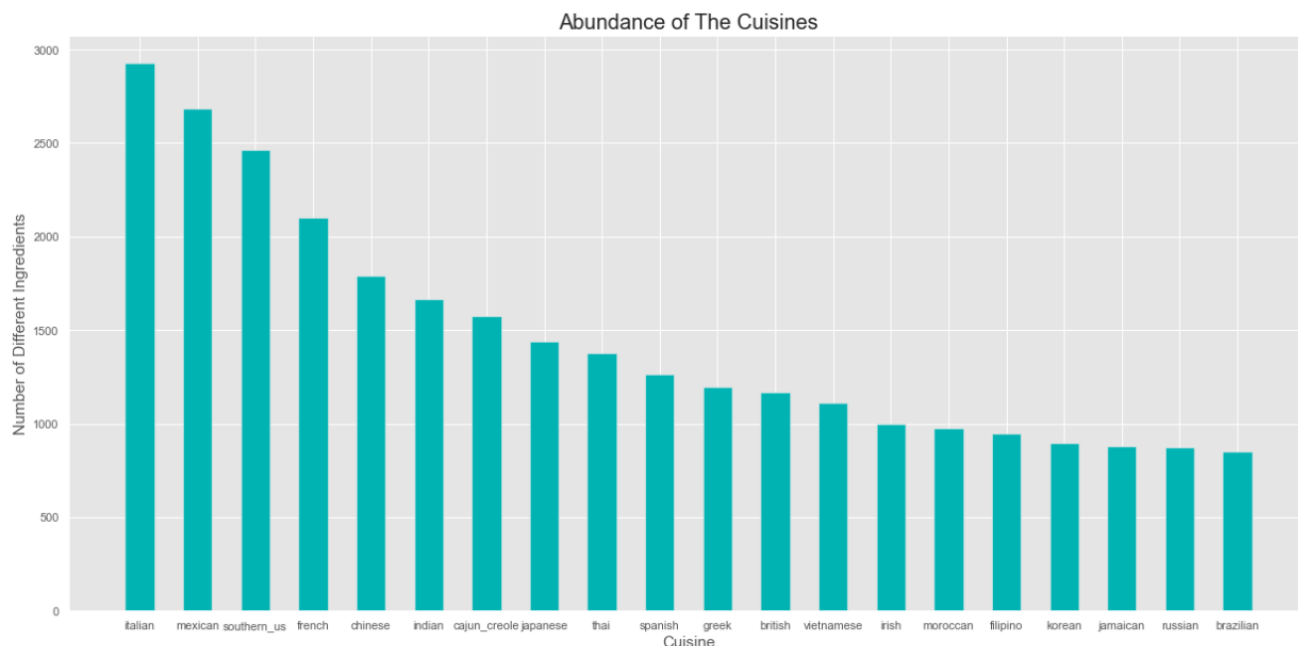


מסקנה מ-2 הגרפים האחרונים

ניתן להסיק מ-2 הגרפים שלעיל כי ישנו קשר מסוים בין קירוב גיאוגרפי של מדינות ומיקומן לבין ממוצע כמות הרכיבים במנות שלהן. למשל:

- כל המטבחים האירופאים בעלי ממוצע כמות הרכיבים במנה אשר קטן מהממוצע הכולל של כמות הרכיבים במנה בכלל המטבחים.
- איטליה וצרפת הממוקמות בסמיכות גיאוגרפית בעלות ממוצעים קרובים של מספר רכיבים במנותיהן (הממוצע של שתיהן נמוך מהממוצע הכללי של כל המדינות).
- תאילנד ווייטנאם אשר ממוקמות גם כן בסמיכות גיאוגרפית בעלות ממוצעים קרובים.
- המטבחים מהמזרח הרחוק (תאילנד, הודו, וייטנאם) בעלי ממוצע מספר רכיבים במנה אשר גבוה יותר מהממוצע הכללי של הרכיבים במנה של כלל המטבחים.

פרט נוסף שעניין אותנו לבדוק היה הגיוון והעושר בסוגי הרכיבים שיש בכל מטבח, על מנת לבחון זאת יצרנו גרף המתאר את כמות סוגי הרכיבים השונים שיש במנות של כל מטבח:



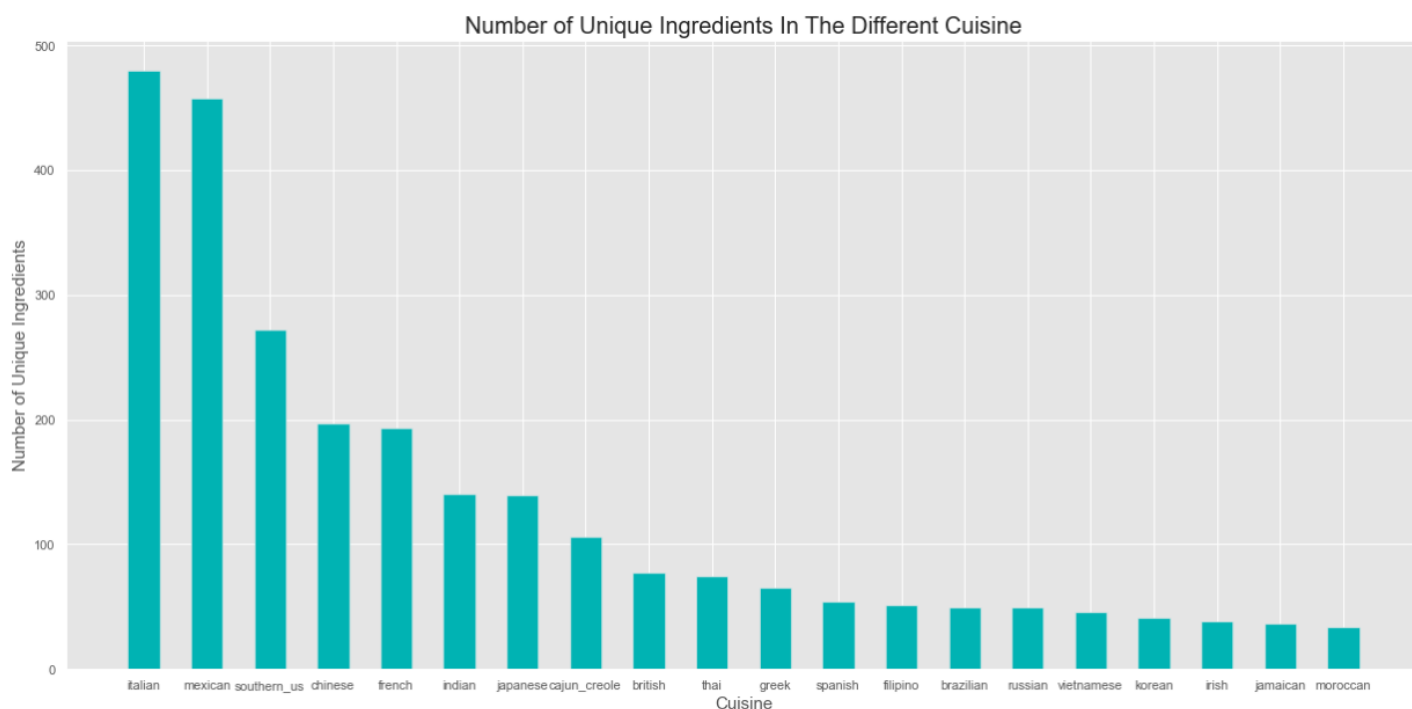
מהגרף הנ"ל ניתן לראות מספר דברים:

- המתכונים בברזיל הכי פחות מגוונים ברכיבים. נזכור כי ב-2 הגרפים הקודמים במטבח הזה במוצע אין המון רכיבים במנה. לא ניתן להסיק מכך דבר שכן כמות המתכונים במאגר הנתונים אשר מיוחס למטבח זה, מועט.
- המטבח האיטלקי בעל הגיוון הרב ביותר, אולם יתכן כי הדבר נובע מכך שבמאגר הנתונים היה למטבח זה את כמות המנות הרב ביותר.

אם כן, להערכתנו לא ניתן להסיק מגרף זה על העושר והגיוון ברכיבים במטבחים השונים שכן עשויה להיות קיימת הלימה מסוימת בין כמות המנות של כל מטבח במאגר הנתונים לבין כמות הרכיבים השונים שיש בו. זאת ראינו ע"י השוואה בין גרף זה לבין הגרף אשר מציג את כמות המנות במאגר הנתונים שיש לכל מטבח: מטבחים עם מספר מנות רב במאגר ימצאו עם גיוון ברכיבים ומטבחים עם מספר מנות מועט להפך.

לאור חוסר יכולת הסקה מהגרף הקודם, ניתחנו את כמות הרכיבים הייחודיים לכל מטבח, כלומר את כמות הרכיבים של כל מטבח אשר קיימים במתכונים שלו בלבד ולא במתכונים של מטבחים אחרים. בכך הנחנו כי נוכל לגזור מסקנה לגבי הרבגוניות שיש בכל מטבח.

להלן הגרף עם כמות הרכיבים הייחודיים לכל מטבח:



על מנת לנתח את הגרף כהלכה השונו גרף זה לגרף המתאר את כמות המנות שיש לכל מטבח, וזאת כדי לא להסיק מסקנה מוטעית אשר נובעת מאי איזון בכמות הנתונים (המתכונים) שהיו על כל מטבח.

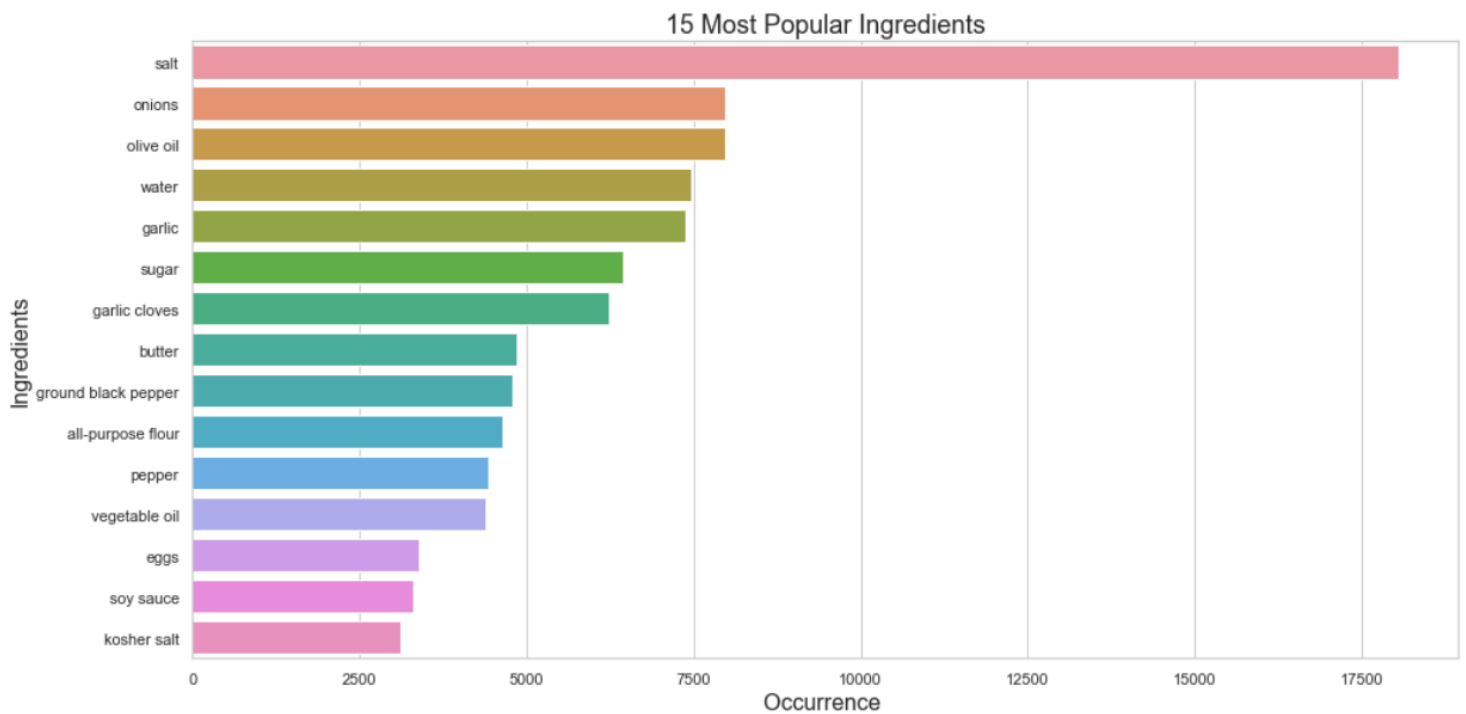
מההשוואה לגרף המתאר את כמות המתכונים הקיימים במאגר בכל מטבח, הבנו כי גם מגרף זה לא ניתן להסיק מסקנות קונקרטיות שכן הממצאים מגרף זה מושפעים מכמות הידע הלא מאוזנת הנתונה על כל מטבח.

בשלב הבא יצרנו ענן עם המיילים של הרכיבים השונים, בחרנו שלא לעשות ניקוי למיילים של רכיבי המתכוננים, שכן יש השפעה רבה של מילות תואר משמעות הרכיבים. להלן ענן המיילים:



מהענן ניתן לראות אילו רכיבים היו עם שכיחות גבוהה בכלל המנות, ניתן לראות כי במטבחים יש שימוש רב ברכיבים הבאים: Purpose Flour, Olive oil, Garlic clove, Black pepper ועוד. הדבר הגיוני שכן מדובר ברכיבי יסוד.

על מנת שנוכל לגזור מסקנות קונקרטיות לגבי פופולאריות הרכיבים במטבחים, יצרנו גרף עמודות המכיל את הכמויות של 15 הרכיבים הפופולאריים ביותר בכלל המטבחים:



מבדיקה זו נראה כי מלח הוא הרכיב הנפוץ ביותר, התוצאה הזו הגיונית שכן מלח הוא חומר גלם נפוץ מאוד בכדור הארץ, כאשר שימושו העיקרי הוא לצורכי מזון. כנ"ל לגבי בצלים, שמן זית, מים ושום אשר מהווים רכיבים בסיסיים נפוצים גם כן, לכן הגיוני כי שכיחותם גבוהה במנות האוכל.

לאחר מכן ביצענו מניפולציות על נתוני הגלם על מנת להציג במערך נתונים חדש ומסודר את 10 הרכיבים הכי פופולאריים של כל מטבח, להלן דוגמא לחלק מטבלת הנתונים החדשה:

moroccan	vietnamese	korean	spanish	greek	japanese	thai	cajun_creole	french	chinese	indian	southern_us	mexican	italian	rate
ground cinnamon	fish sauce	soy sauce	olive oil	olive oil	soy sauce	fish sauce	green bell pepper	sugar	soy sauce	water	sugar	avocado	olive oil	0
olive oil	sugar	sugar	sugar	pepper	sugar	soy sauce	water	unsalted butter	sugar	tomatoes	eggs	tomatoes	water	1
ground ginger	water	water	water	tomatoes	water	sugar	olive oil	olive oil	water	tumeric	pepper	black beans	pepper	2
tumeric	soy sauce	sesame seeds	kosher salt	ground black pepper	sake	water	chicken broth	water	eggs	garam masala	water	pepper	sugar	3
water	kosher salt	brown sugar	tomatoes	feta cheese	eggs	lime juice	dried thyme	ground black pepper	brown sugar	curry powder	milk	olive oil	ground black pepper	4
sugar	lemongrass	eggs	eggs	fresh dill	mirin	lime	andouille sausage	eggs	dark soy sauce	clove	butter	water	eggs	5
tomatoes	red chili peppers	fresh ginger	large eggs	water	dashi	brown sugar	celery ribs	butter	fresh ginger	sugar	large eggs	kosher salt	fresh basil	6
saffron threads	lime juice	black pepper	pepper	sugar	brown sugar	unsweetened coconut milk	pepper	pepper	light soy sauce	salt	unsalted butter	shredded cheddar cheese	grated parmesan cheese	7
kosher salt	chiles	green onions	ground black pepper	eggs	sesame seeds	kaffir lime leaves	ground black pepper	large eggs	low sodium soy sauce	curry leaves	kosher salt	jalapeno chilies	kosher salt	8
pepper	eggs	minced garlic	black pepper	black pepper	salt	lemongrass	black pepper	kosher salt	white pepper	fresh ginger	ground black pepper	lime juice	black pepper	9

מסקנות

מבחינה של 10 הרכיבים הכי פופולאריים בכל מטבח גילינו את הפרטים הבאים:

- המלח הוא הרכיב הכי פופולארי במטבחים הבאים (15 מטבחים מתוך 20): יוון, דרום ארה"ב, פיליפינים, הודו, ג'מייקה, ספרד, איטליה, מקסיקו, בריטניה, קייג'ון, ברזיל, צרפת, אירלנד, מרוקו, רוסיה.
- רוטב הסויה הוא הרכיב הכי פופולארי ב-3 מטבחים: סין, יפן וקוריאה.
- רוטב דגים הוא הרכיב הכי פופולארי ב-2 מטבחים: תאילנד, ווייטנאם.

רכיבים ייחודיים לאזור:

- הגרם מסאלה רכיב פופולארי בהודו, הוא הרכיב שלישי הכי שכיח במטבח זה.
- בקוריאה הסומסום שכיח במאכלים, שמן סומסום הוא הרכיב השלישי הכי שכיח במטבח זה, וכן זרעי סומסום הינם הרכיב השמיני הכי פופולארי במטבח זה.
- במטבח היפני יש 3 רכיבי יסוד מבוססים אורז בעלי שכיחות גבוהה במנות: מירין הינו הרכיב השלישי הכי שכיח, הסאקה ממוקם במקום הששי והחומץ אורז הינו הרכיב ה-7 הכי שכיח במטבח זה.
- הקשאסה, המשקה הברזילאי הלאומי, מהווה אחד מ-10 הכי פופולאריים במנות, והוא ממוקם במקום ה-8.
- תבלין קייג'וני הינו רכיב פופולארי (מקום 8) במטבח הקייג'וני.

*בנספח קיימים סרטוטים (גרפים וענן מילים) אשר יצרנו עבור מטבחים שונים.

סיכום תובנות

ראינו שאוכל מושפע ממקומות גיאוגרפיים שונים ומתרבויות שונות ואנו רואות צורך ללמידה מעמיקה יותר של מאכלים וקשריהם למאפיינים של חברה. הדבר דרוש על מנת להבין בני אדם ותרבויות ועל הדמיון באוכל שהם חולקים.

צעדים להמשך

האתגר הראשוני להמשך הינו לאמן מודל machine learning על מנת לחזות על סמך רכיבים של מנה את שייכותה הגיאוגרפית.

רעיון נוסף לפיתוח הינו בניית מודל נוסף של machine learning שקולט מיקום גיאוגרפי ופולט רשימת רכיבים פוטנציאליים עימם ניתן להכין מנות מסוימות באותו אזור. המודל הזה ילמד את ההעדפות המשתמש ובכך יאמן את יכולתו להמליץ למשתמש על רכיבים שהוא יאהב.

נושא נוסף שהיינו רוצות לבחון הינו הקשר בין המטבחים של המדינות לבין רווחה, בריאות ואושר.

אתגרים

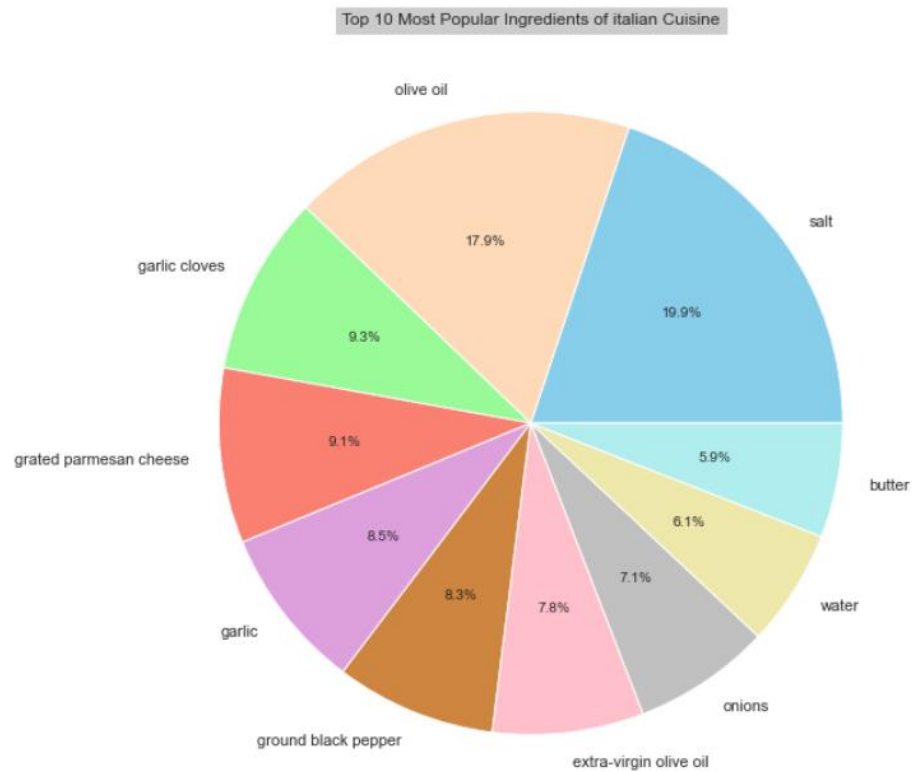
האתגרים המרכזיים בעבודה היו שונים, חלקם נבעו מבעיות ומקשיים בבניית הקוד, חלקם נבעו מסנכרון בין המטרות והעבודה שלנו כצוות במקביל בתקופה של ריחוק חברתי. כמו כן האתגרים נבעו מהצורך המתמיד להיות ביקורתיים כלפי העבודה של עצמנו והממצאים שעולים במיוחד בכל הנוגע לנכונות הנתונים שאנו מקבלות ולמסקנות שניתן לגזור מהם.

סיכום

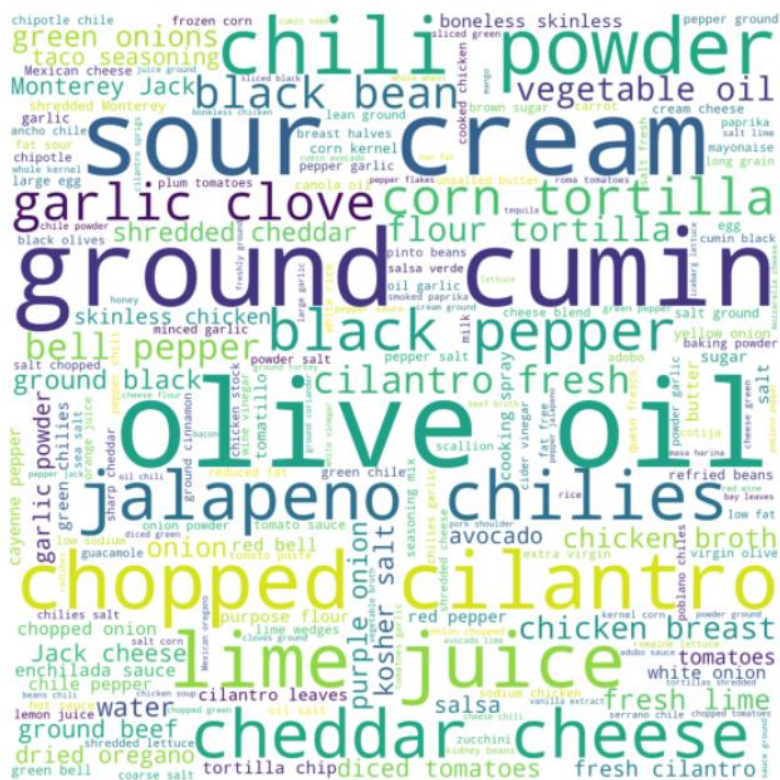
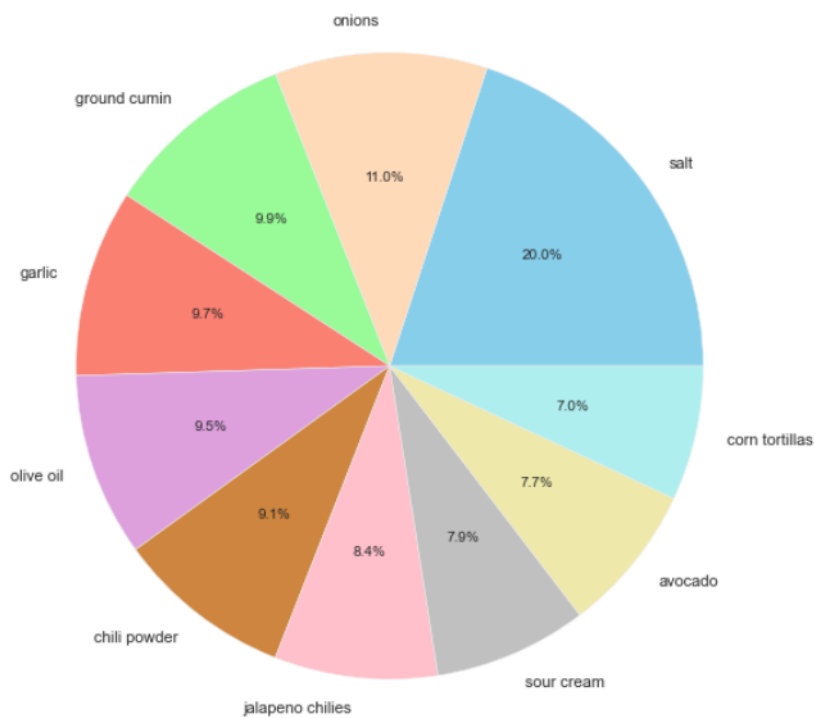
לסיכום, פרויקט זה היה מעניין ומאתגר עבורנו. למדנו ושיפרנו את יכולותינו בכלים הטכנולוגיים בהם התעסקנו לאורך הסמסטר. נדרשנו לעבוד ביחד, לחלק עבודה ולהיות מתואמות. פיתחנו רעיון מאפס ועד לביצוע סופי. בעינינו, נושא זה מעניין ונרחב ויש עוד המון מה לחקור ולגלות בתחום. מחקרים עתידיים יכולים להעמיק ולדייק את הקוד המשווה בין המדינות השונות.



איטליה

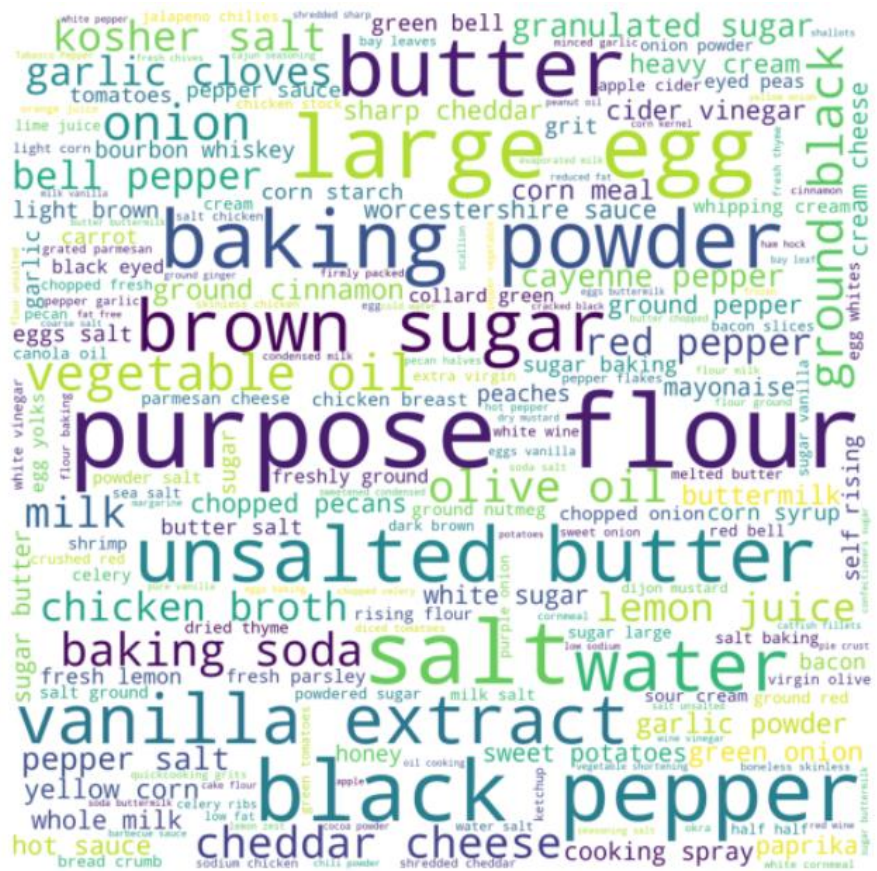
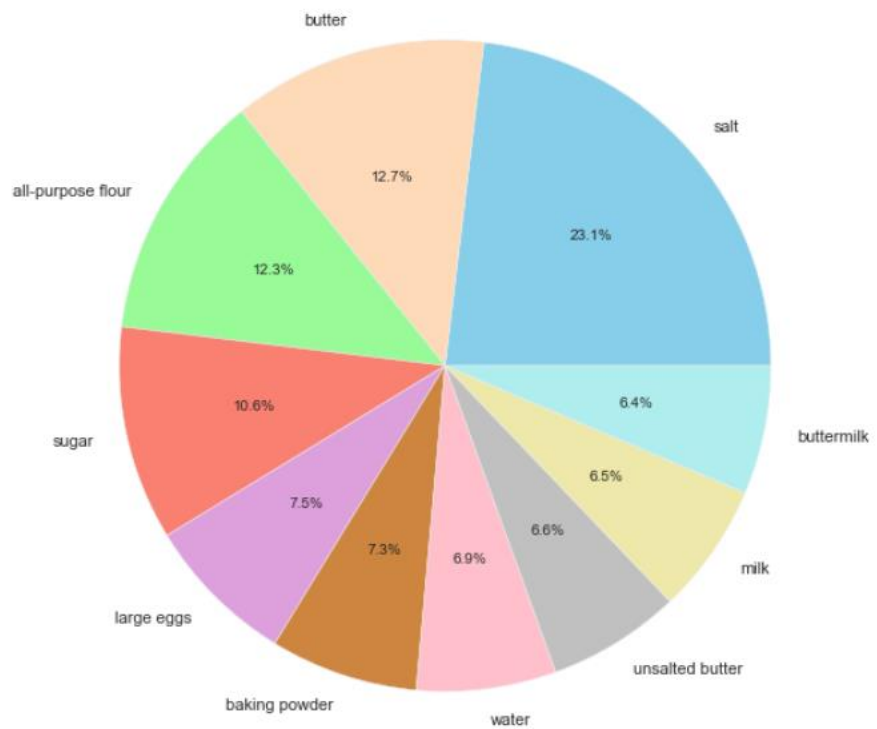


Top 10 Most Popular Ingredients of mexican Cuisine

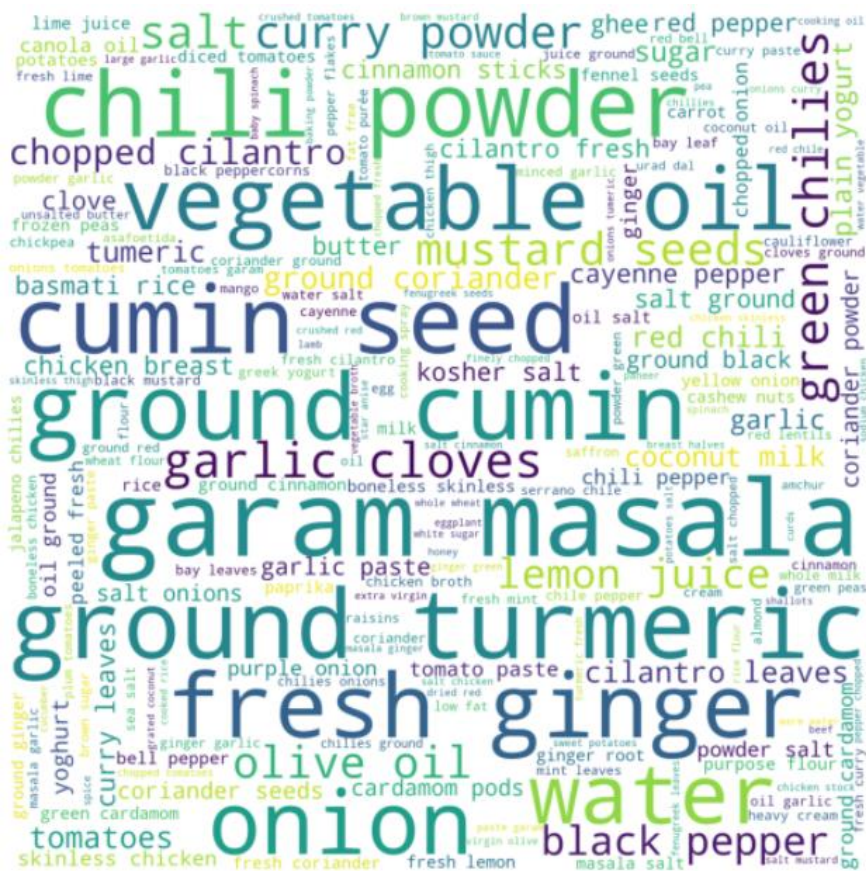
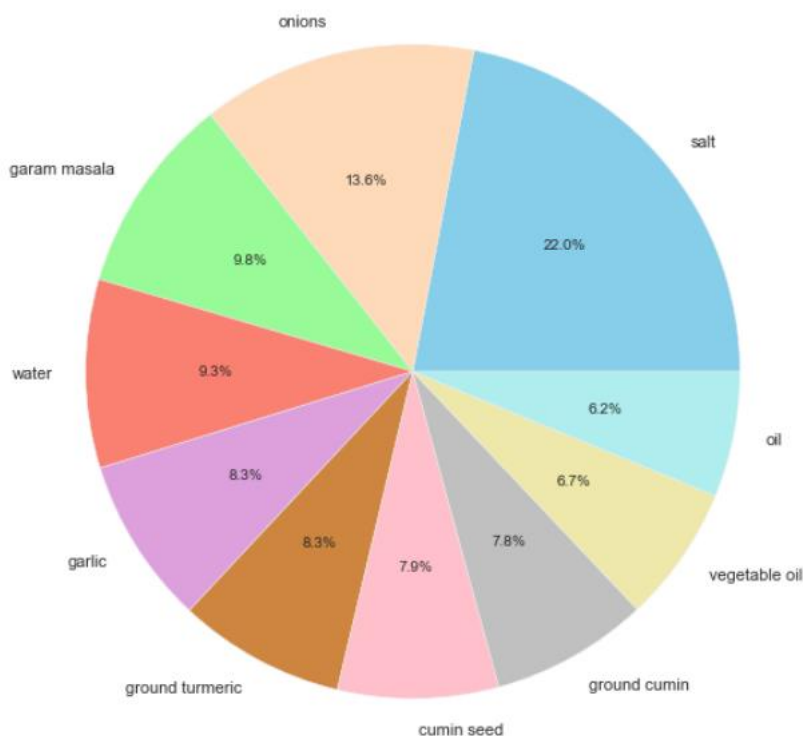


דרום ארה"ב

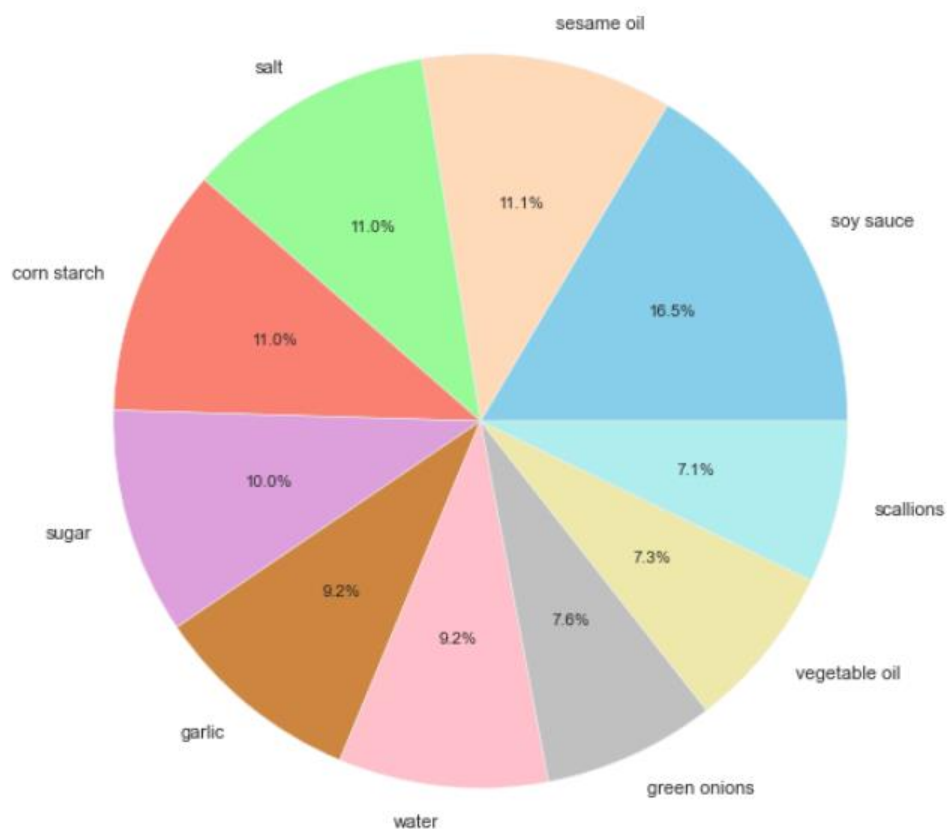
Top 10 Most Popular Ingredients of southern_us Cuisine



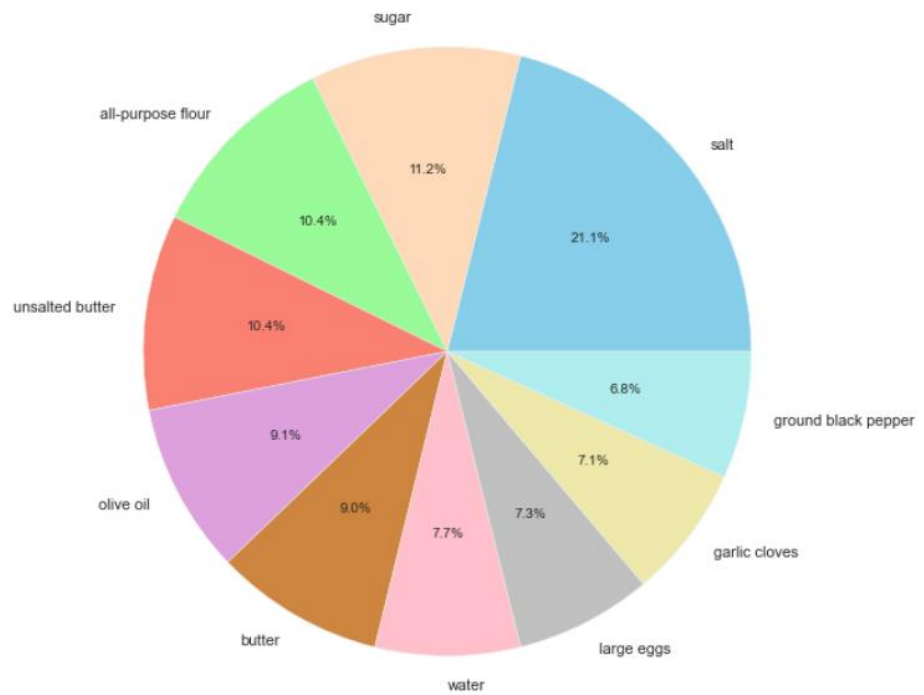
Top 10 Most Popular Ingredients of indian Cuisine



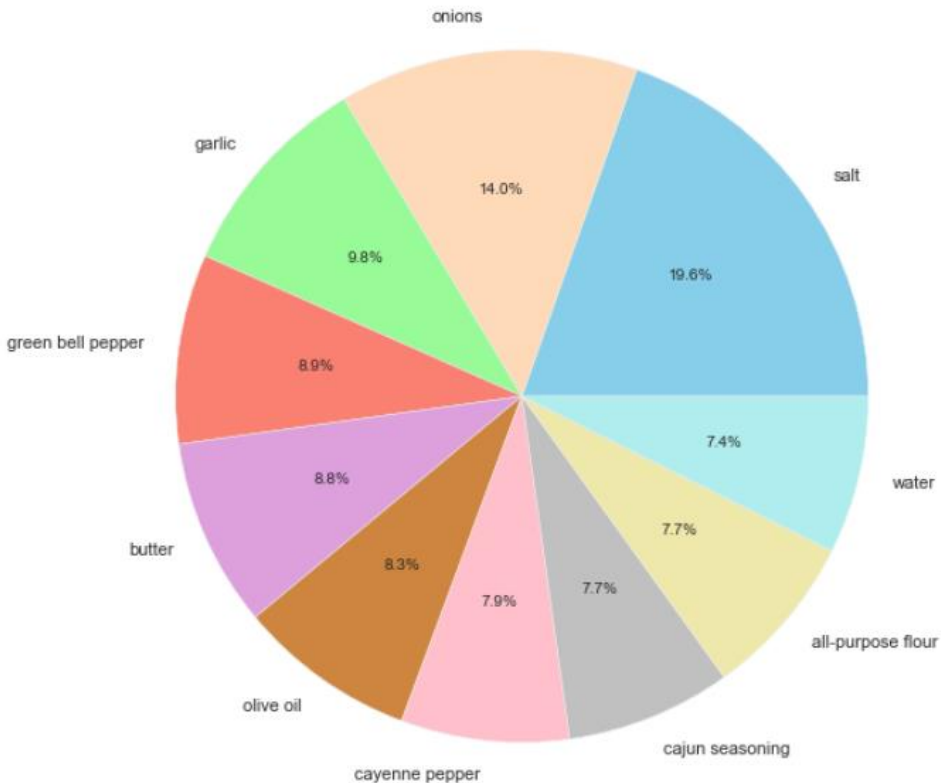
Top 10 Most Popular Ingredients of chinese Cuisine



Top 10 Most Popular Ingredients of french Cuisine

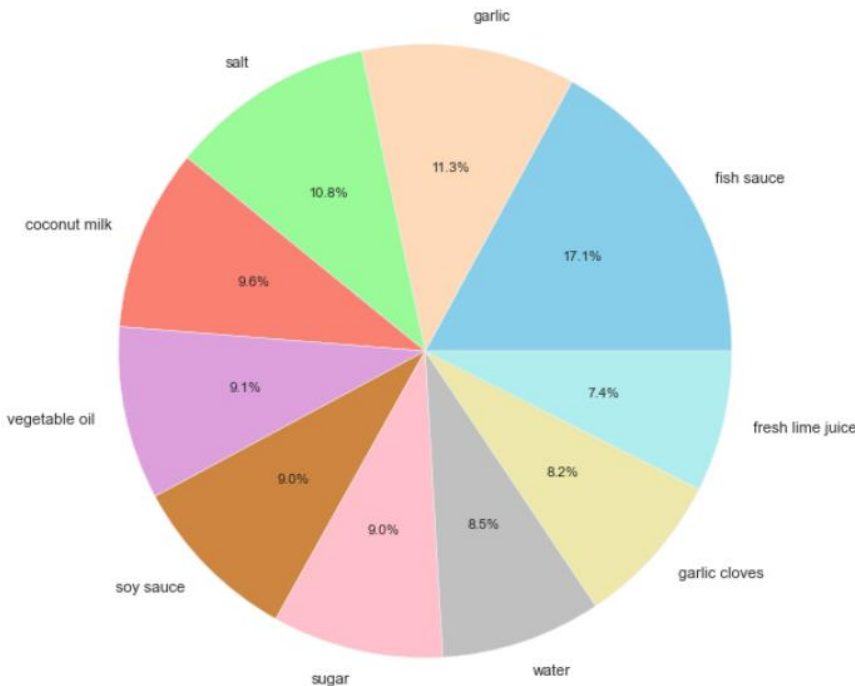


Top 10 Most Popular Ingredients of cajun_creole Cuisine

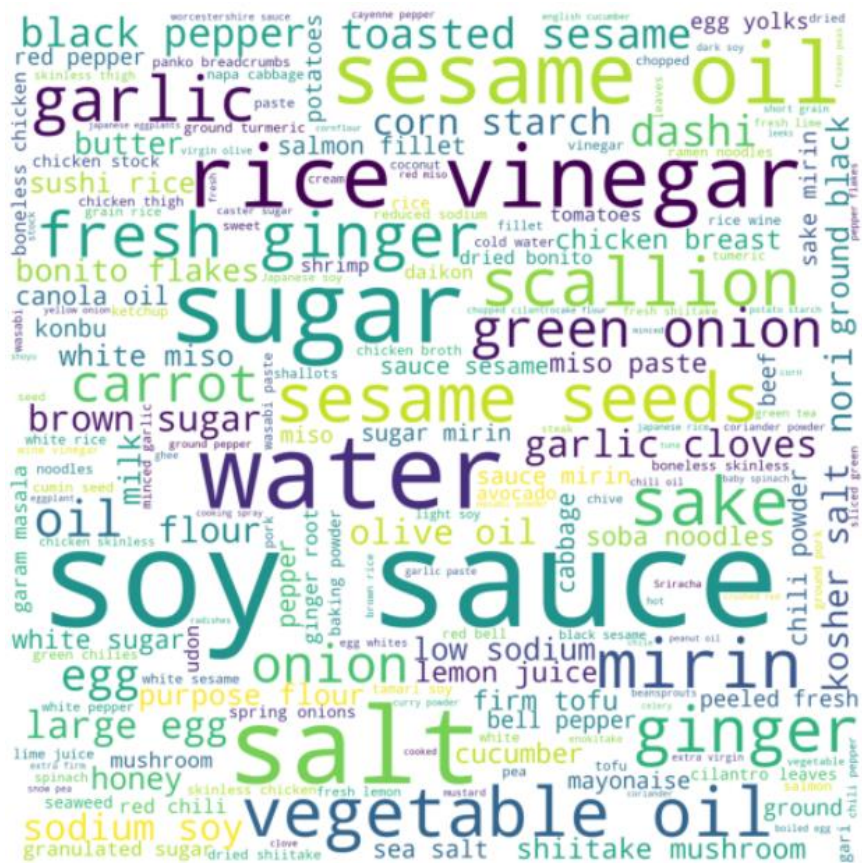
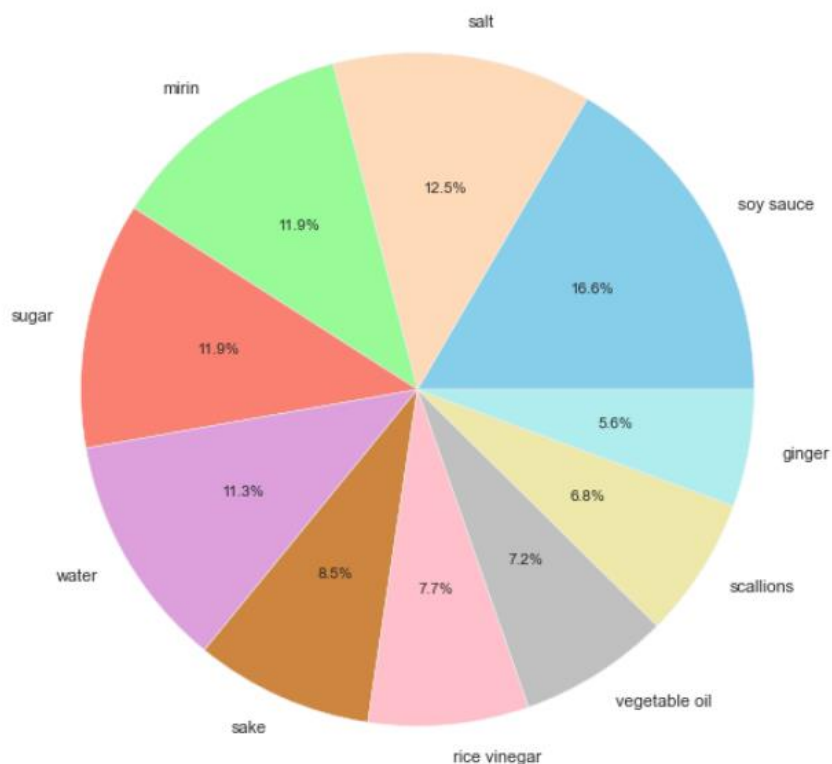


תאילנד

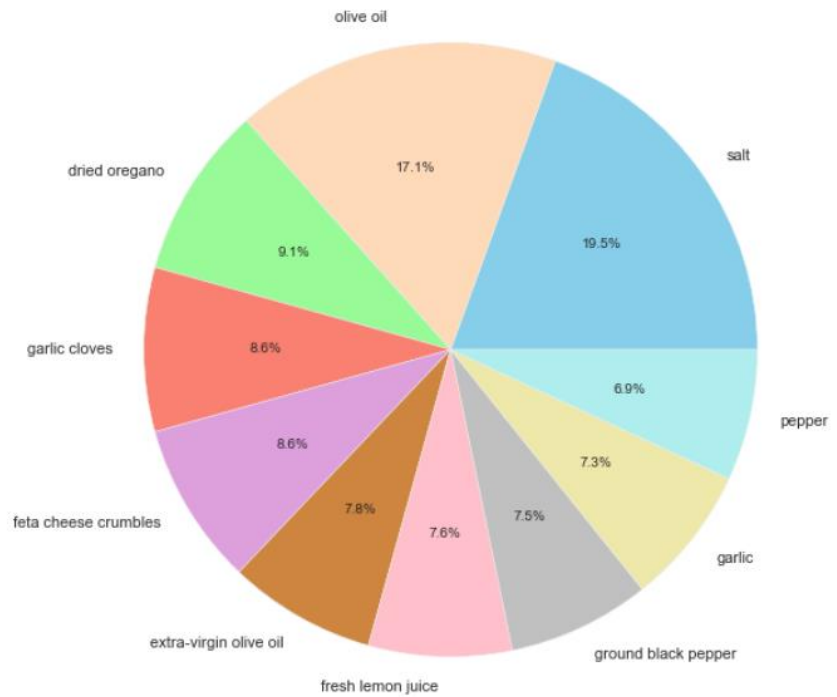
Top 10 Most Popular Ingredients of thai Cuisine



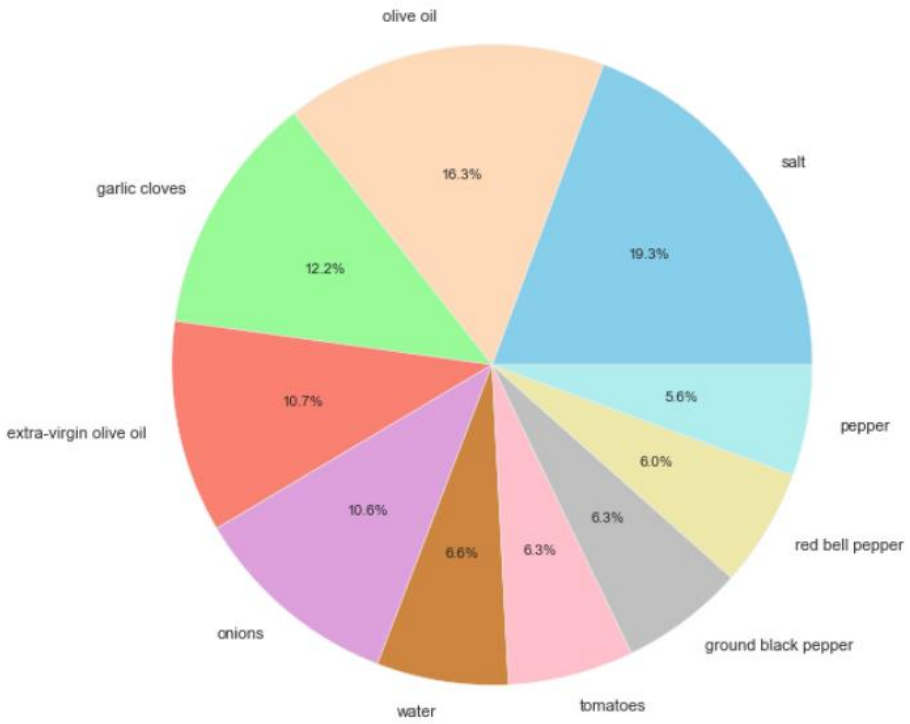
Top 10 Most Popular Ingredients of Japanese Cuisine



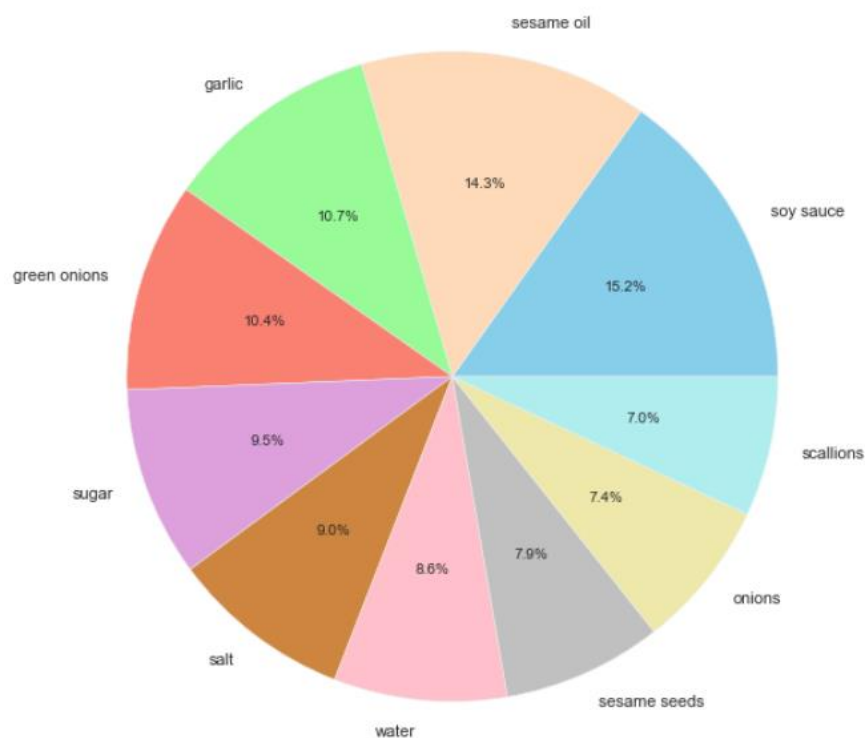
Top 10 Most Popular Ingredients of greek Cuisine



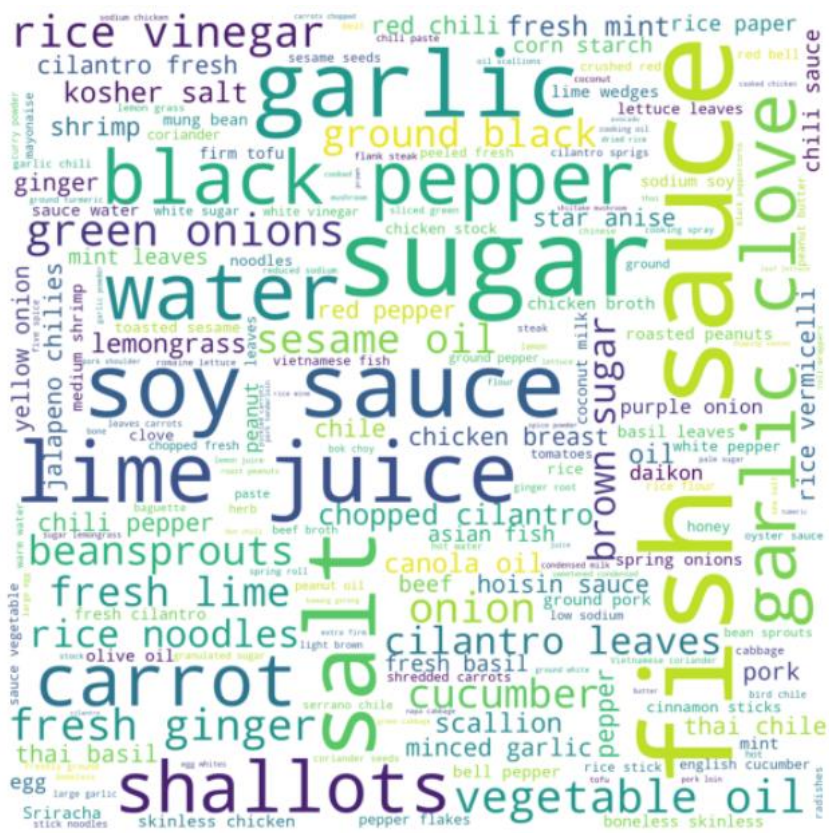
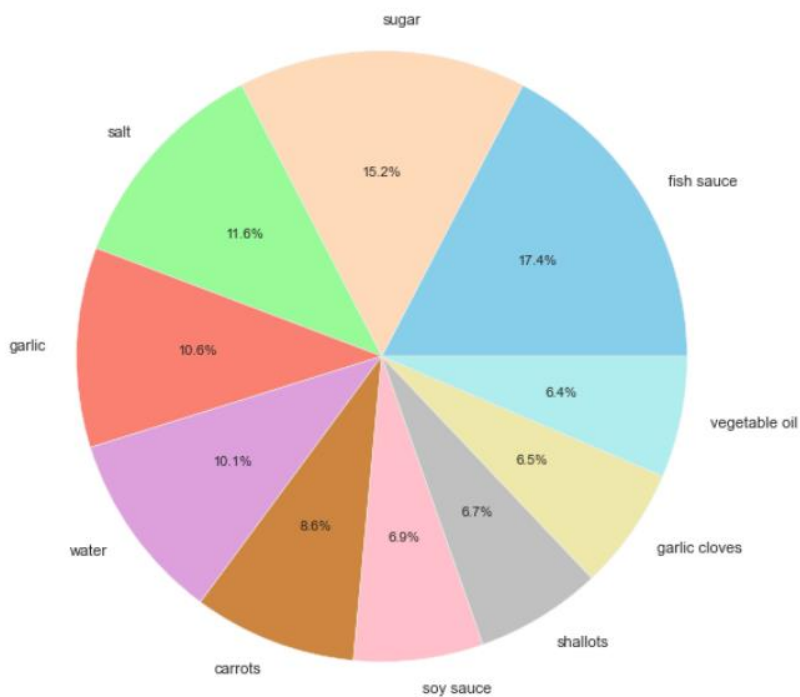
Top 10 Most Popular Ingredients of spanish Cuisine



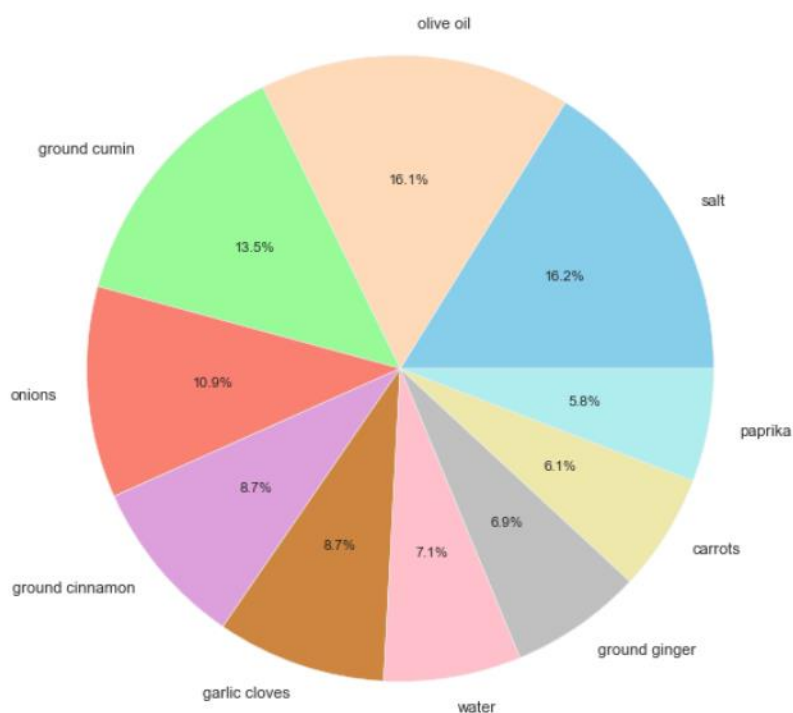
Top 10 Most Popular Ingredients of korean Cuisine



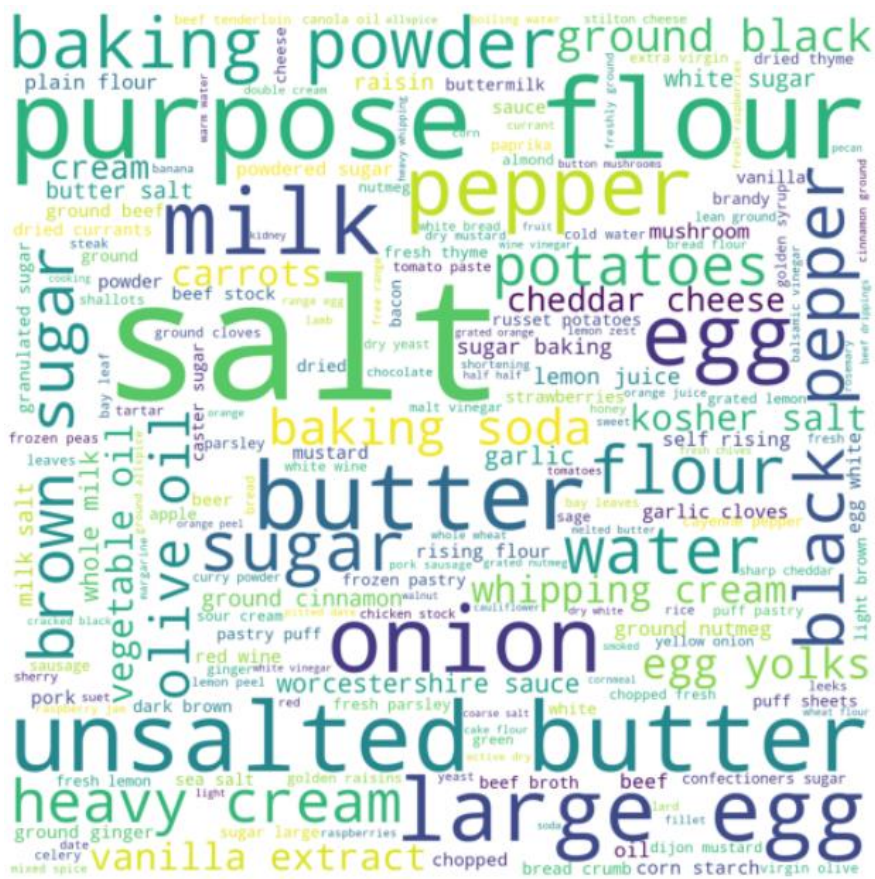
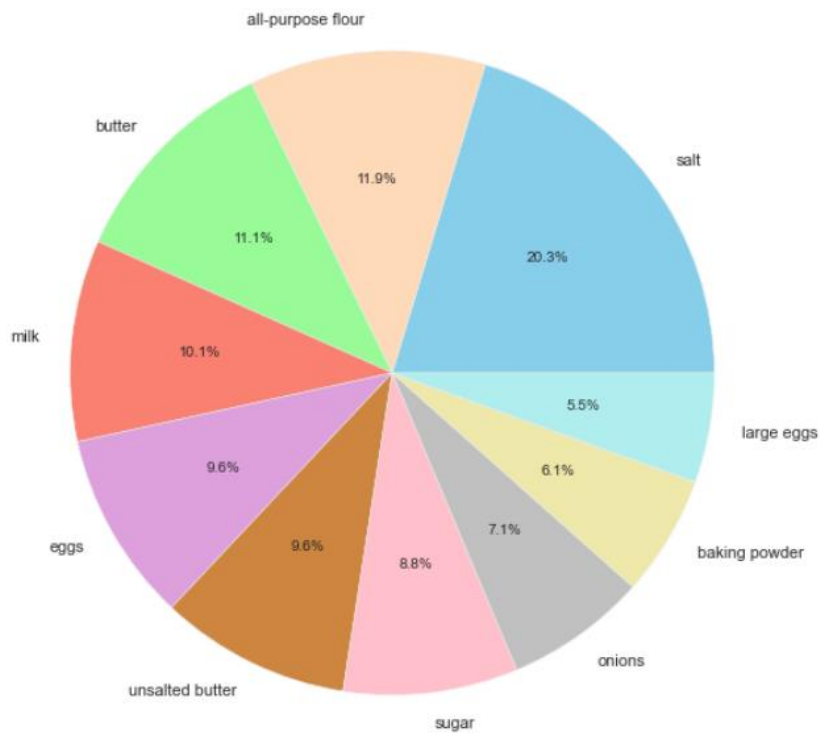
Top 10 Most Popular Ingredients of vietnamese Cuisine



Top 10 Most Popular Ingredients of moroccan Cuisine

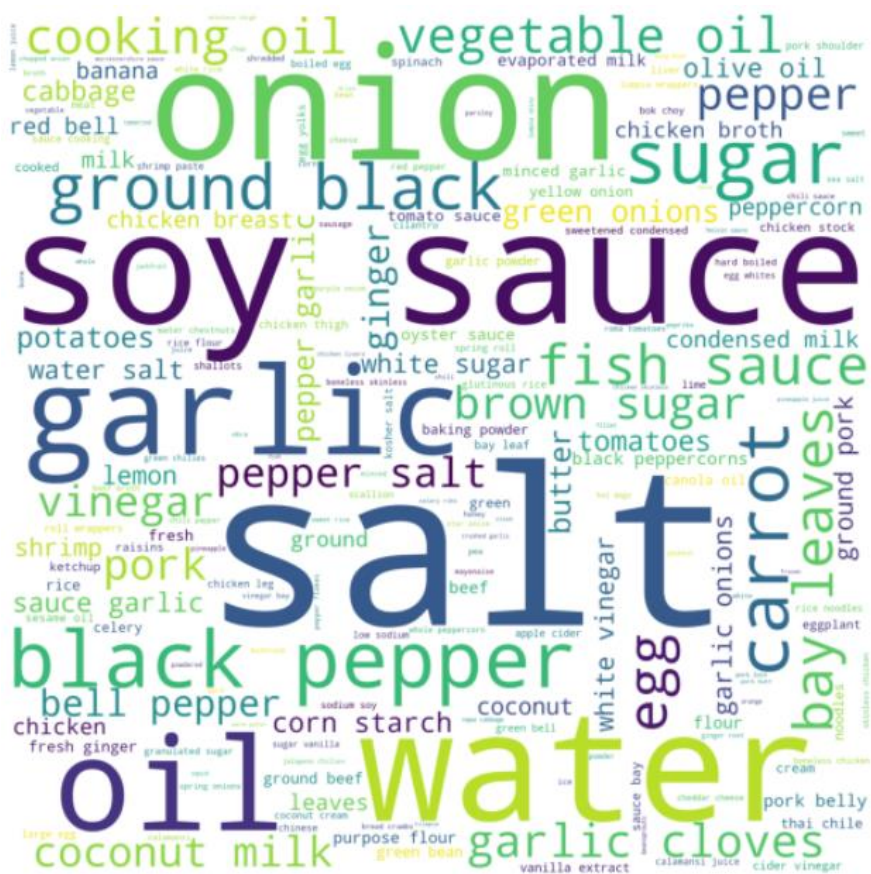
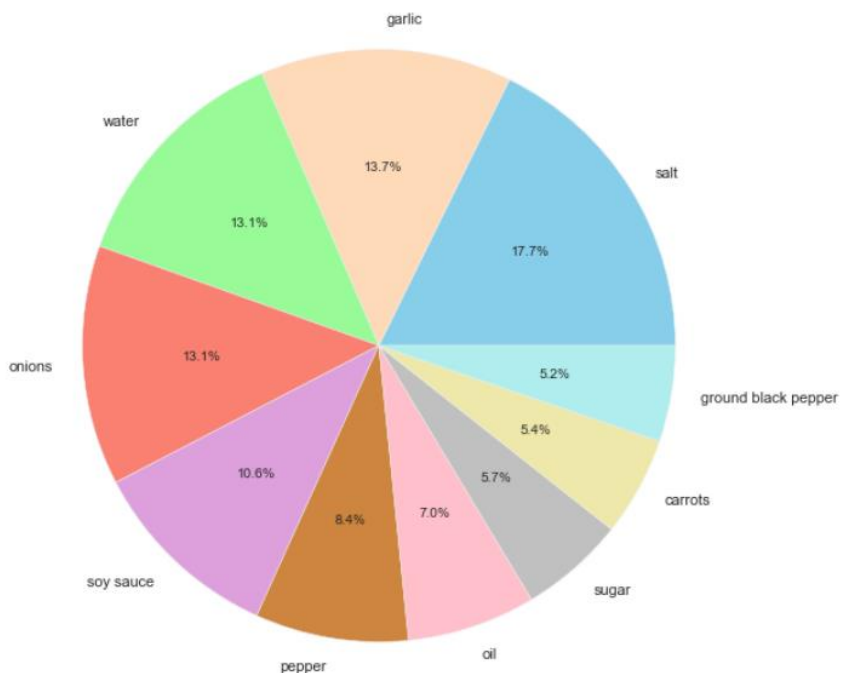


Top 10 Most Popular Ingredients of british Cuisine

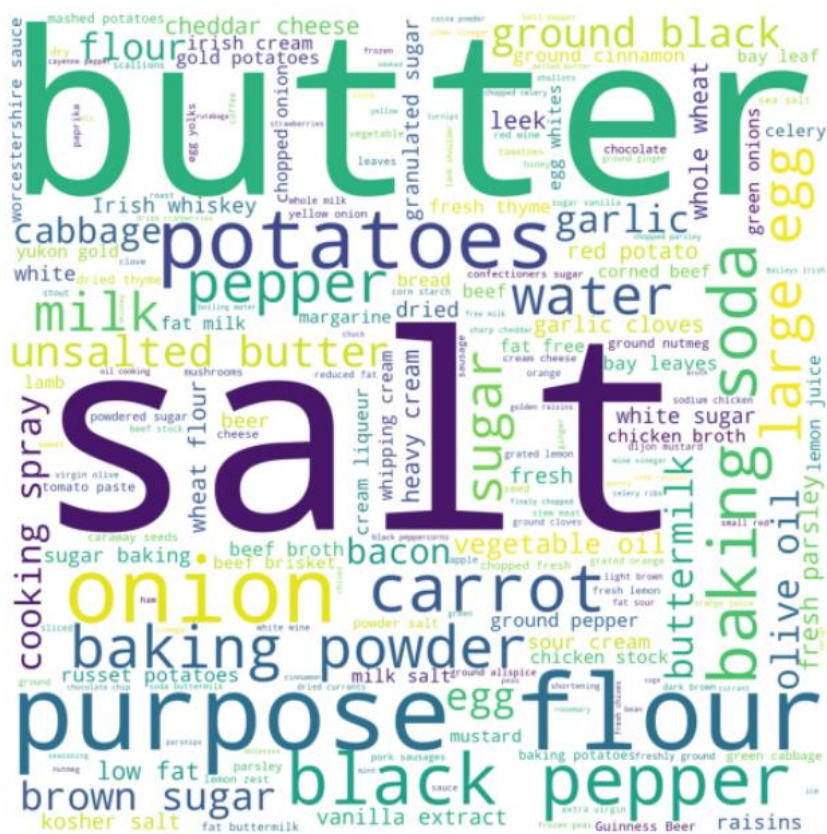
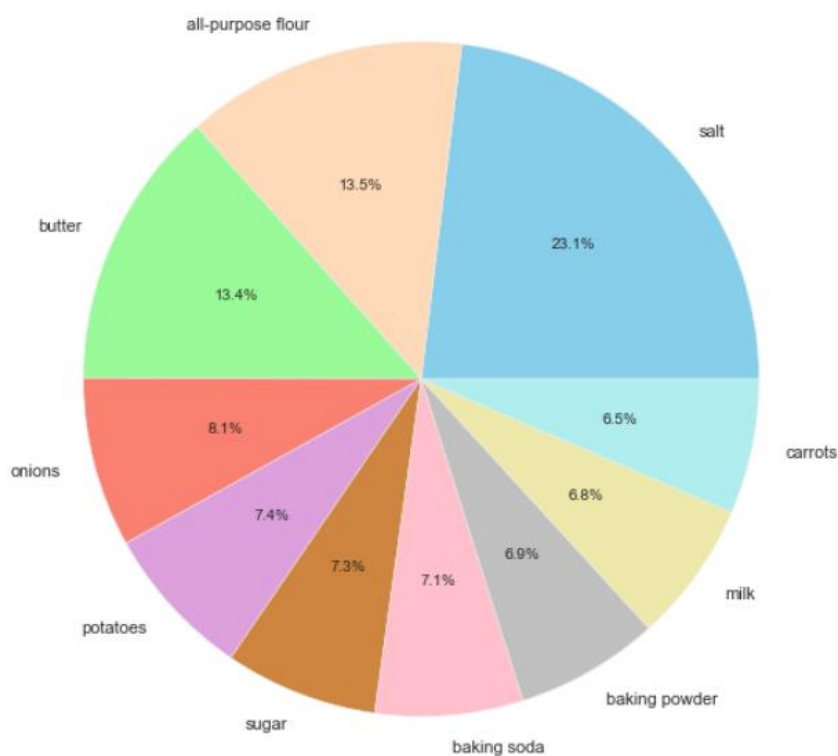


פיליפינים

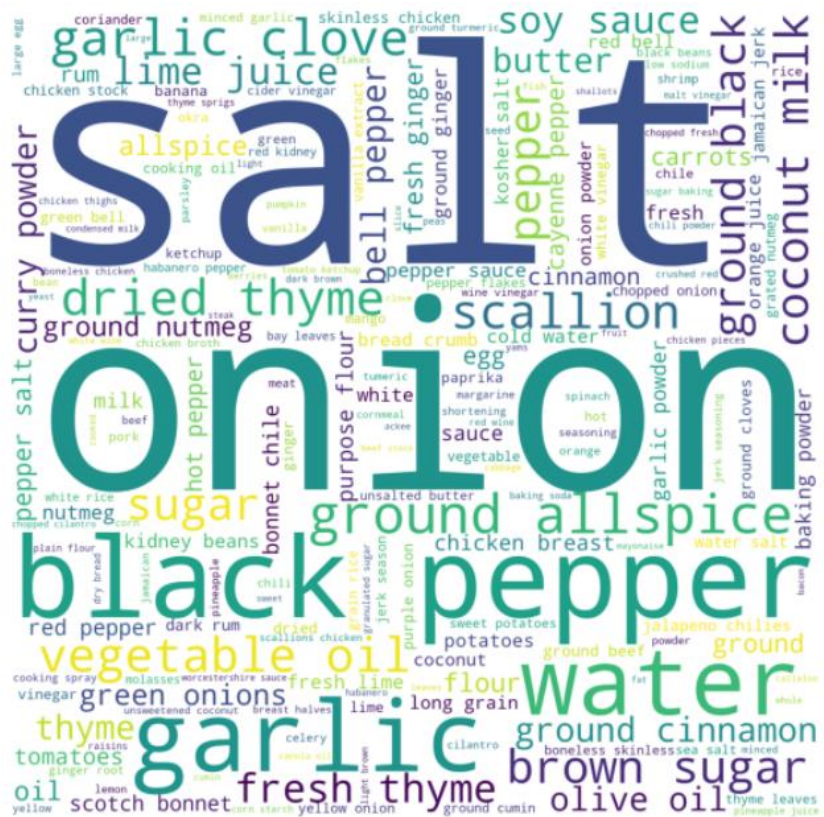
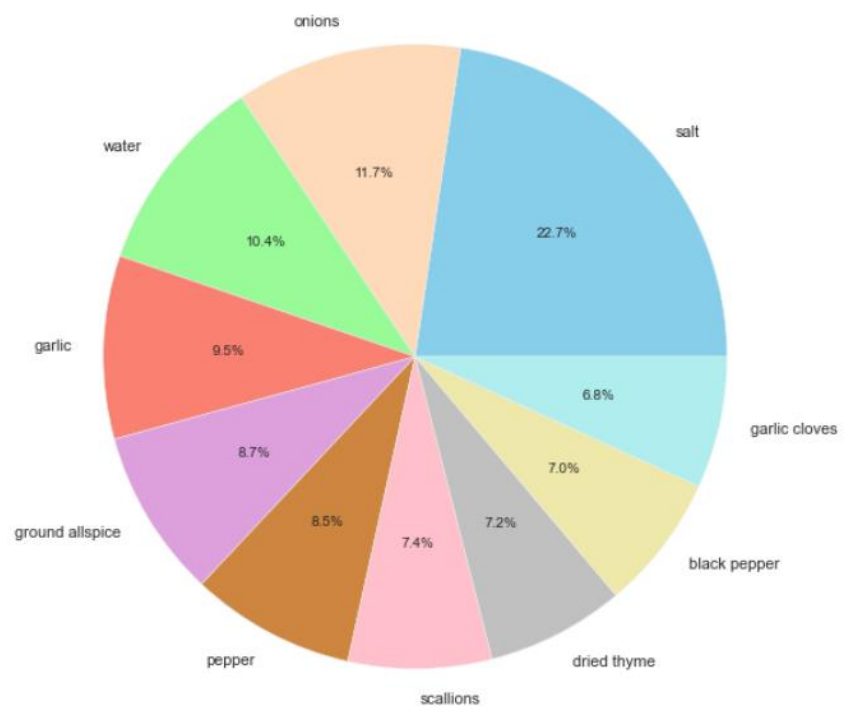
Top 10 Most Popular Ingredients of filipino Cuisine



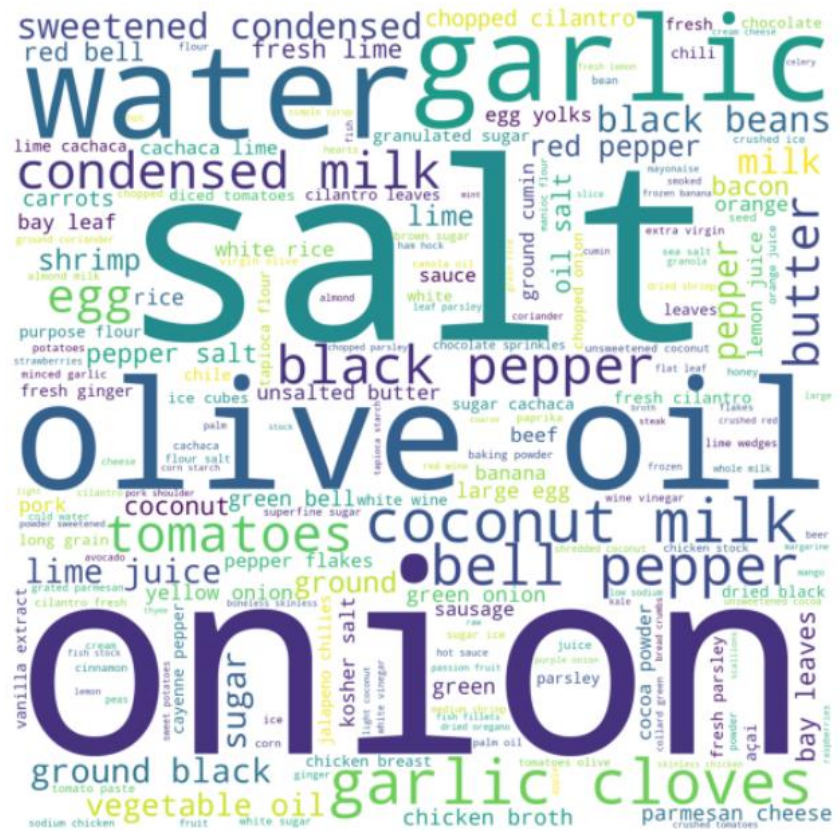
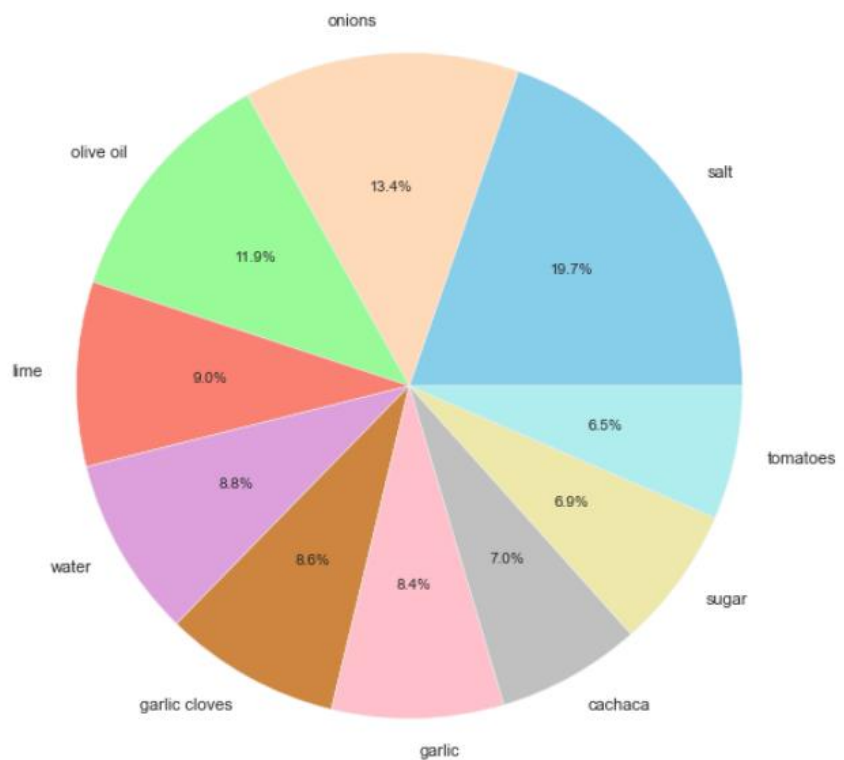
Top 10 Most Popular Ingredients of Irish Cuisine



Top 10 Most Popular Ingredients of jamaican Cuisine



Top 10 Most Popular Ingredients of brazilian Cuisine



Top 10 Most Popular Ingredients of russian Cuisine

