LAB MANUAL
PART A
(PART A: TO BE REFFERED BY STUDENTS)

Experiment No-10

A.1 Aim:
Develop social media text analytics models for improving existing product/ service by analyzing customer's reviews/comments

Lab Objective	To design and develop social media analytics models.
Lab Outcome	Design and develop content and structure based social media analytics models. Design and implement social media analytics applications for business

A-2 Prerequisite

A.3 OutCome
Students will be able to collect, monitor, store and track social media data.

A.4 Theory:
Google Trends tells us what people are searching for, in real time. We can use this data to measure search interest in a particular topic, in a particular place, and at a particular time.

1. Selecting a topic
2. Reading the interest over time graph
3. Understanding the numbers
4. Searches by location
5. Top and rising searches
6. Data excluded from search

Selecting a topic

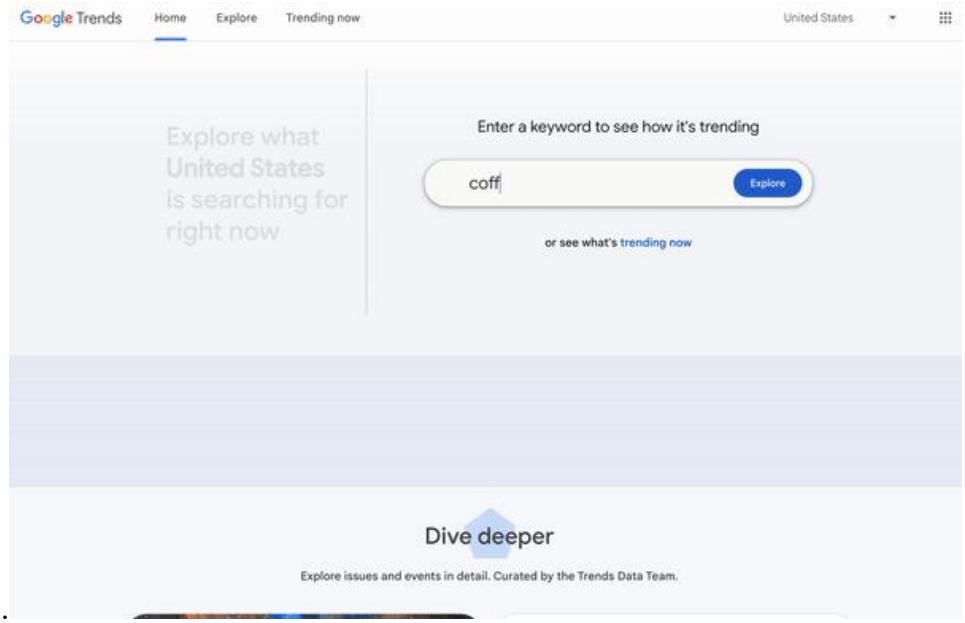
You can explore Google Trends data using the free Trends Explore tool. The data is grouped into topics to give you the fullest picture of what people are searching for. You can also query individual search terms.

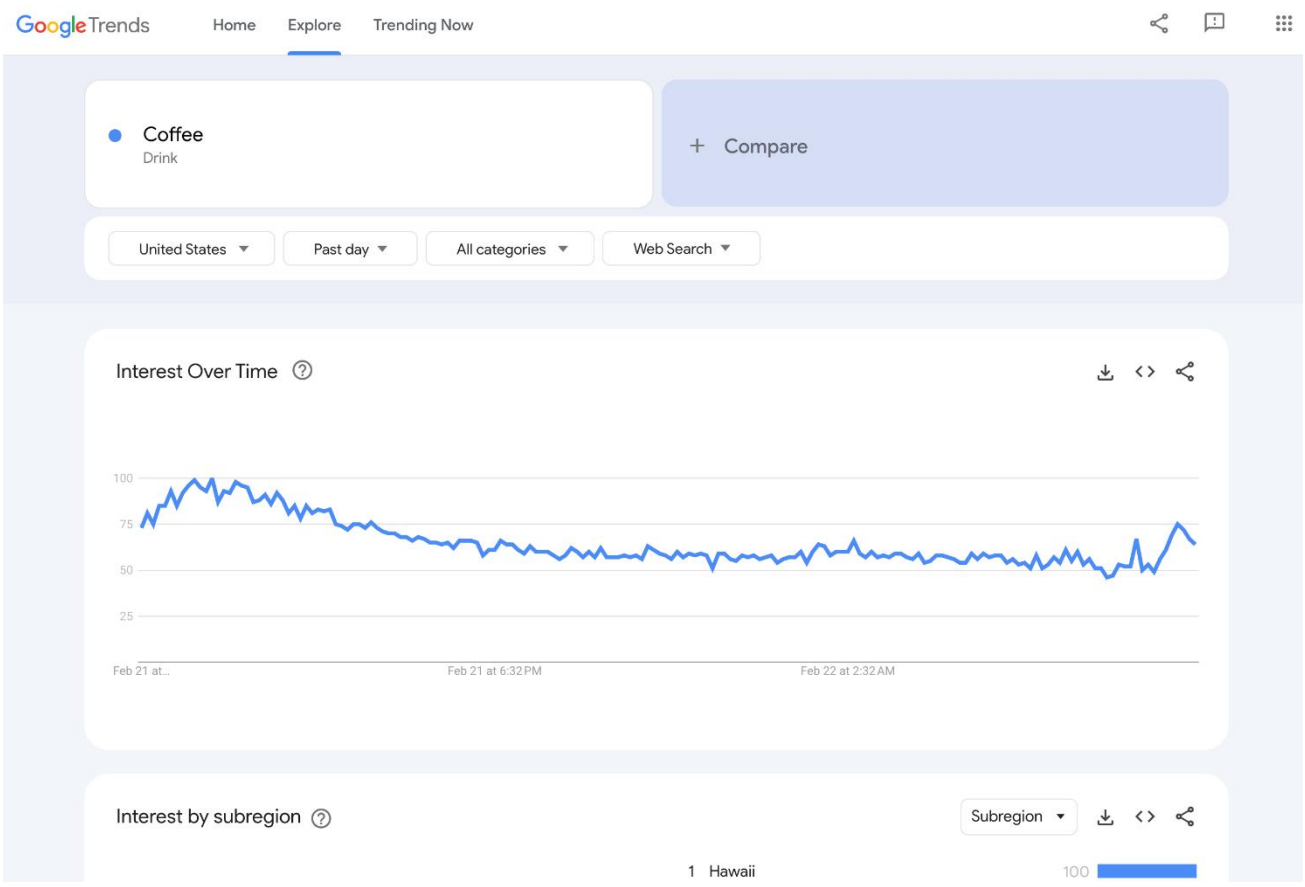
STEP 1
As you type your query, a drop-down list will appear. At the top is the exact search term you typed. Below, you'll see a list of topics which Google Trends believes you may be trying to research.

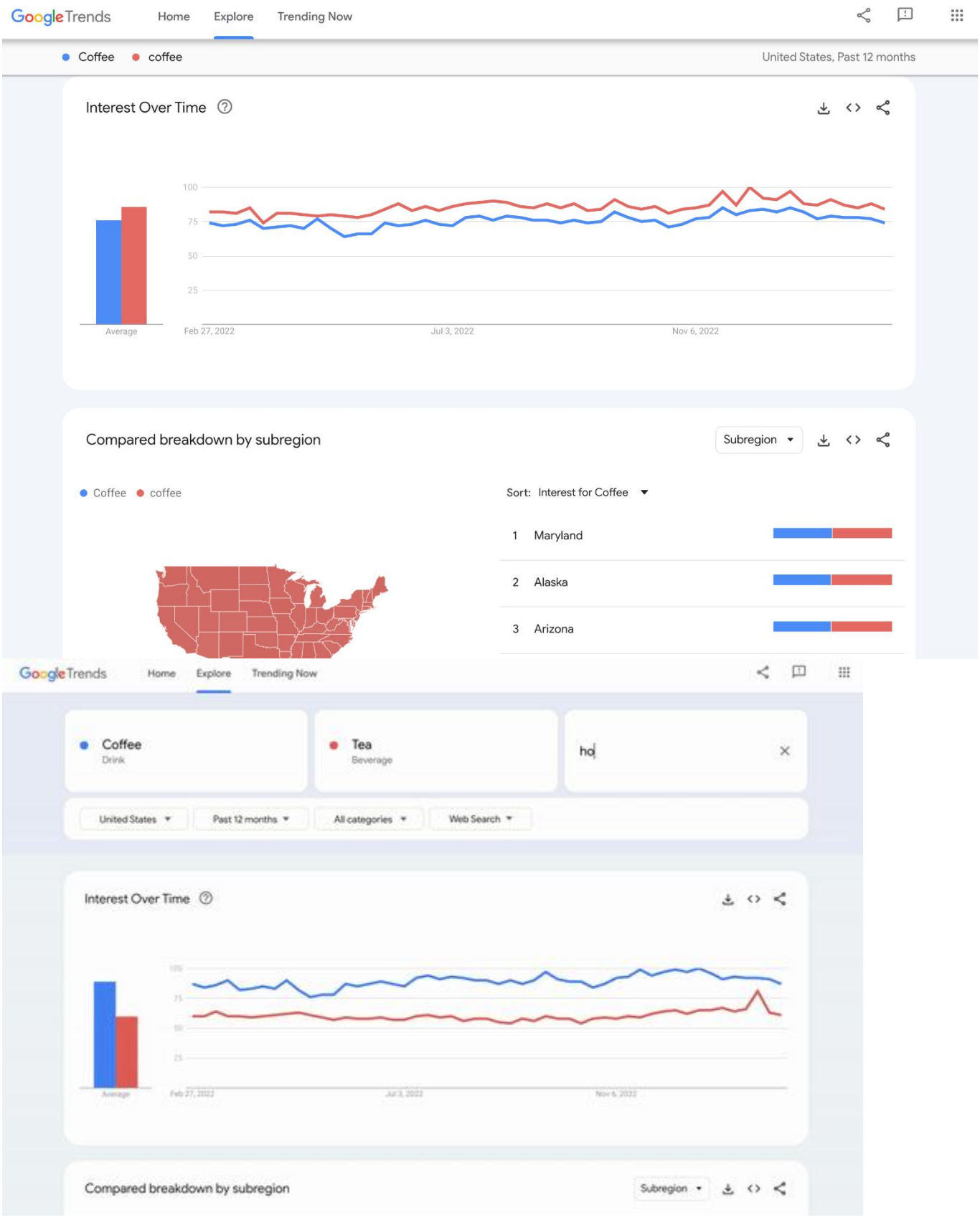
STEP 2
If you see the relevant topic in the drop-down list, click on it to view the data. Otherwise, use the search term.Topics are generally considered to be more reliable for Google Trends data. They pull in the exact phrase as well as misspellings and acronyms, and cover all languages. This is more useful, particularly when looking at world data.

STEP 3
To get a better understanding of the difference between topics and search terms, explore a search term and a topic with the same title, at the same time. You will notice the results look different.

STEP 4
You can explore up to five topics or terms on the Trends Explore tool. Make sure you compare like for like -- topics should be compared with other topics, and search terms should be compared with other search terms.







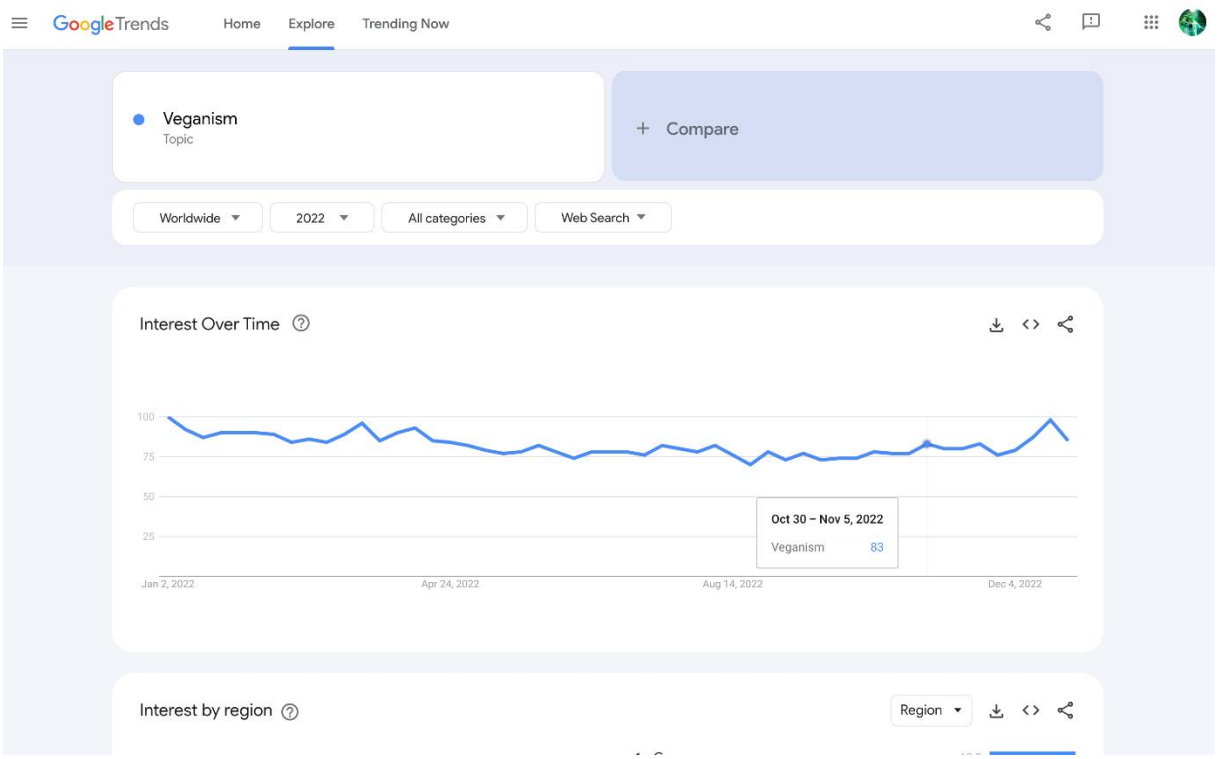
Reading the interest over time graph

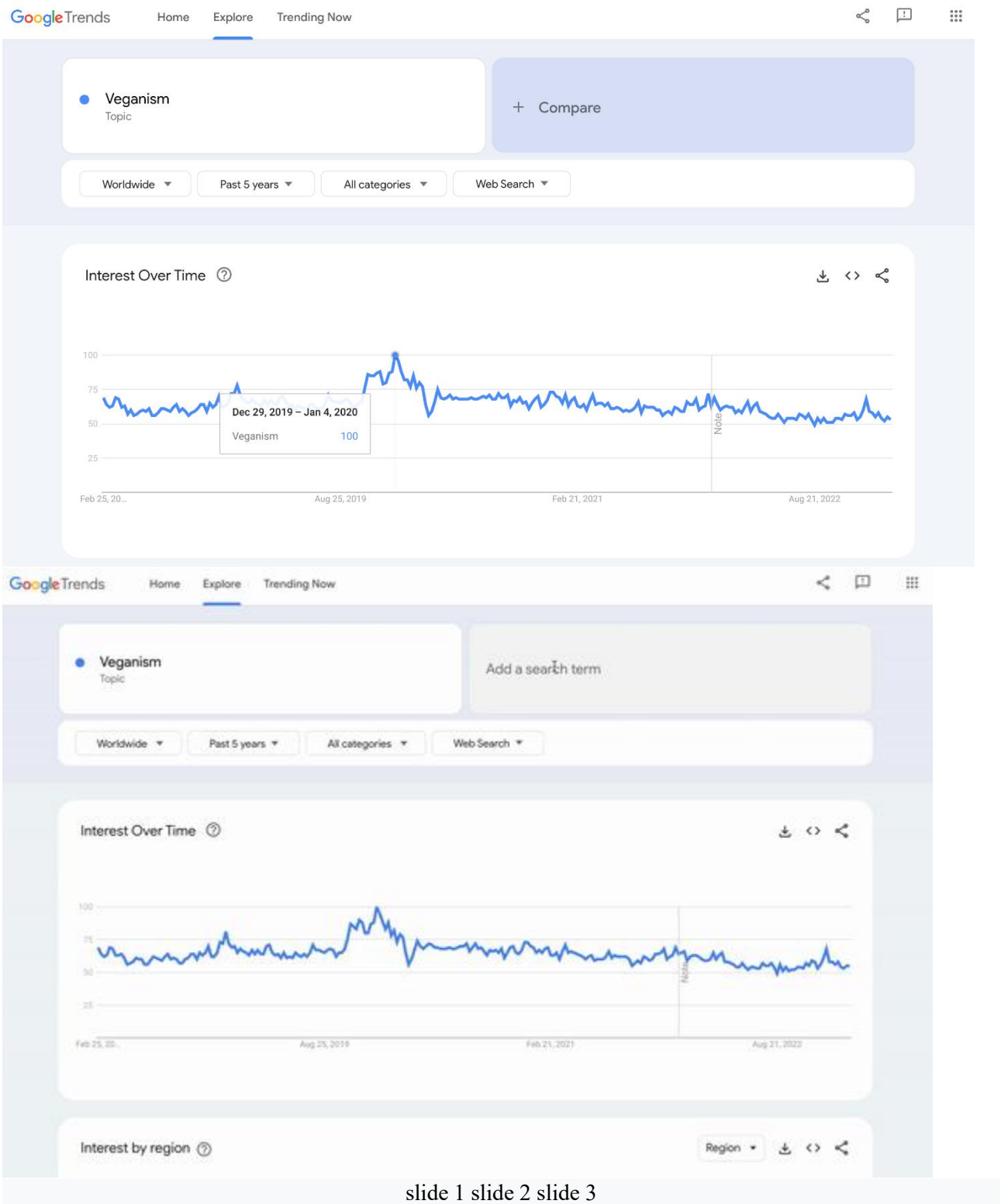
When you search for a topic on Trends Explore, you'll see a graph showing its popularity over time. Being able to read this graph is the first step to understanding the numbers behind Google Trends.

STEP 1
Let’s search the topic Veganism. Use the drop-down menus to specify Worldwide and the year 2017. Hover over the graph to see search interest over time, numbered from 1-100. Find the maximum value on the graph and make a note of the date.

STEP 2
Let’s change the time range of the data to Past 5 years.The maximum value on the graph is still 100, but the date of this value has changed. See if you can locate the peak from the previous graph, too.

STEP 3
The values on the graph do not represent absolute search volume. Instead they are normalized, then indexed on a scale from 1-100. Each point on the graph is divided by the highest point, 100.See how the maximum points change again when we click + Add comparison and type in another topic, such as ketogenic diet.





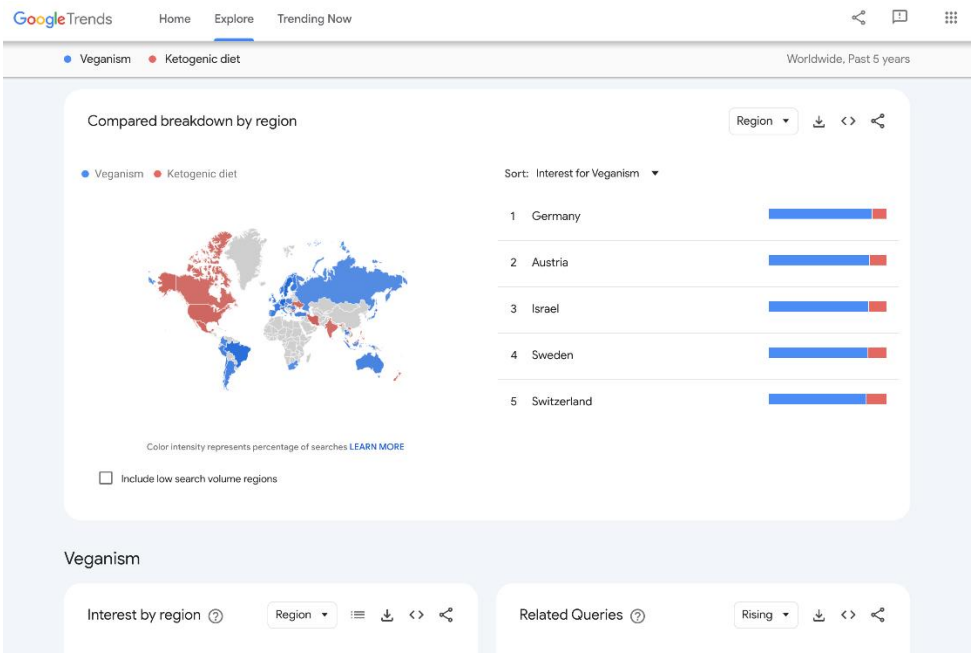
slide 1 slide 2 slide 3

Understanding the numbers

By now, we've already heard that data on the Trends Explore tool is indexed and normalized. But what does this actually mean?

Indexing: Google Trends data is pulled from a random, unbiased sample of Google searches, which means we don't have exact numbers for any terms or topics. In order to give a value to terms, we index data from 1-100, where 100 is the maximum search interest for the time and location selected.

Normalization: When we look at search interest in a topic or query, we are not looking at the total number of searches. Instead, we look at the percentage of searches for that topic, as a proportion of all searches at that time and location.



Searches by location

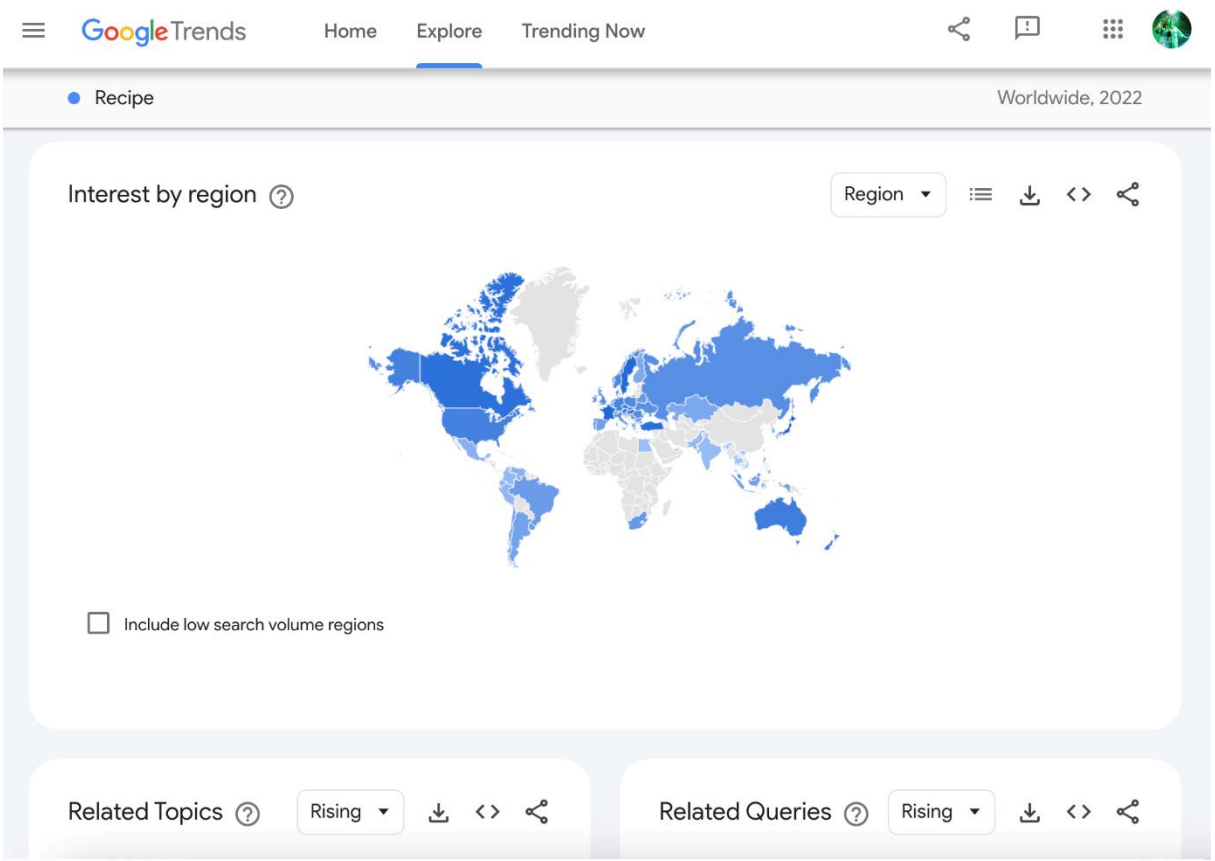
Google Trends also breaks down data by location. In this section, we will explore the basics of location data with Google Trends, and understand the numbers behind this.

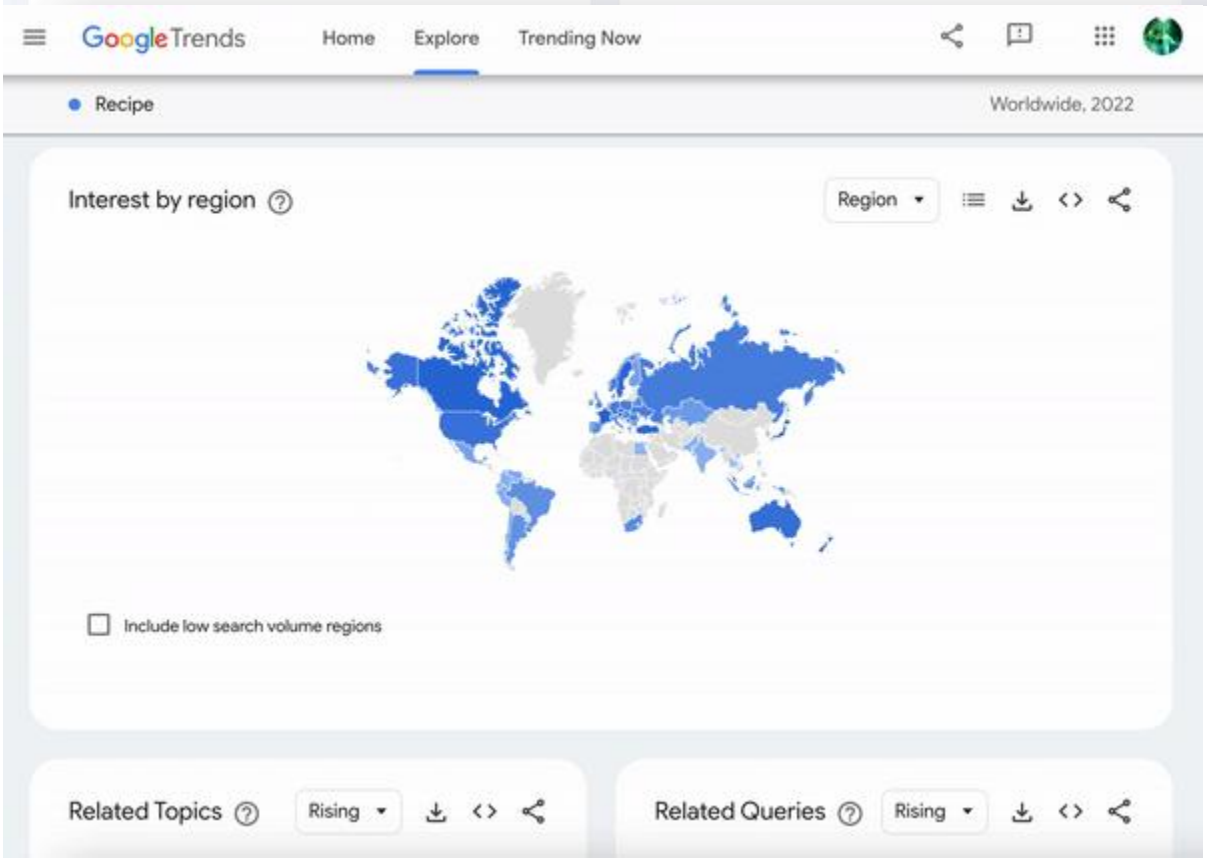
STEP 1
Once you have performed your Google Trends search, scroll down to see search interest by location.If your location is set to Worldwide, you will see search interest by country. If you have selected a country, you will see Interest by region. Darker-shaded areas have a higher level of search interest.

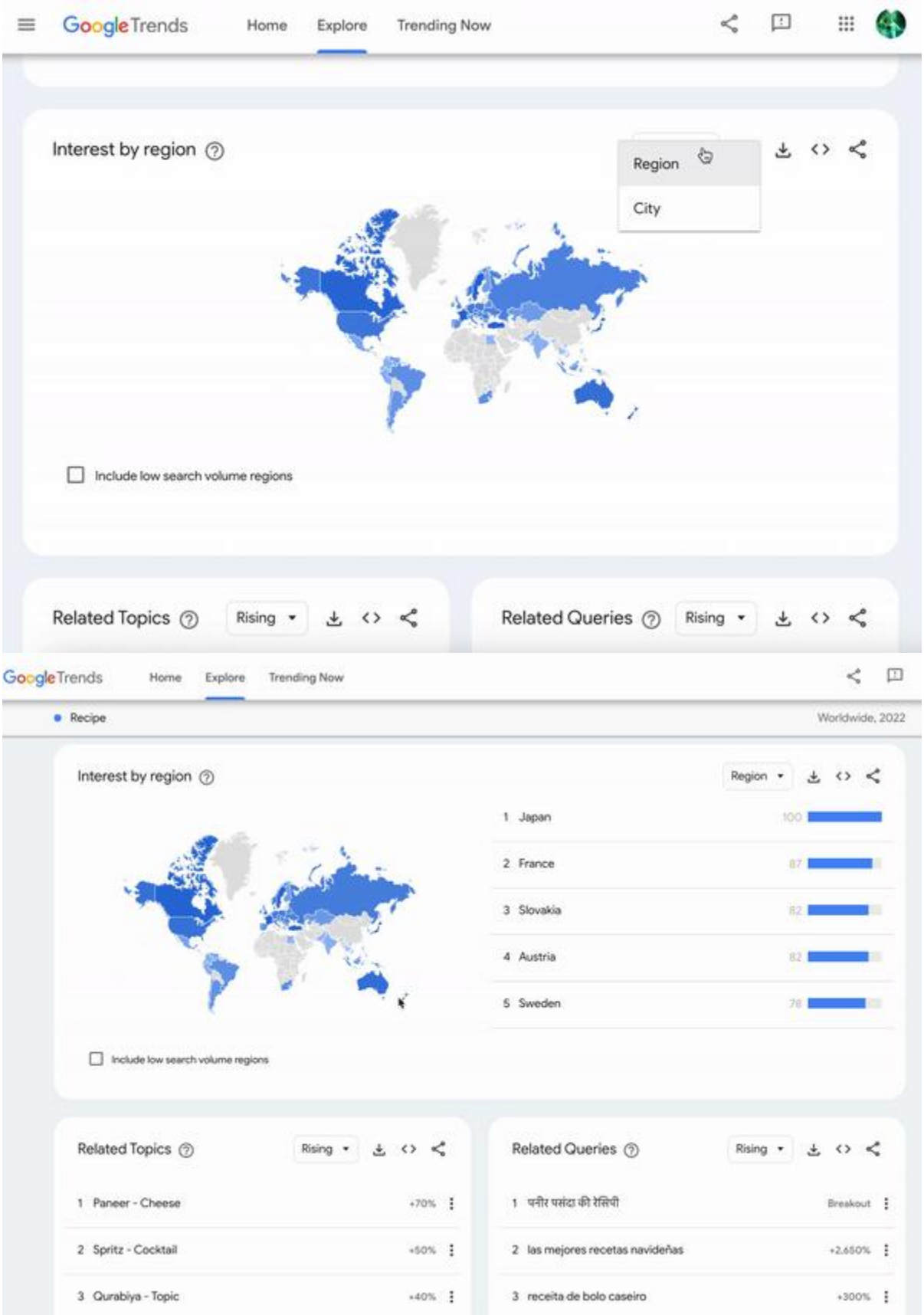
STEP 2
Trends Explore will automatically show you the map view of your data. Toggle between map view and list view using the button with three stacked horizontal lines, at the top right of the map. If your window is wide enough, you will be shown both views simultaneously.

STEP 3
You can view Google Trends data grouped by country, region, and city. In some countries, you can also view search interest by metro area. To select the breakdown you want, use the menu at the top right of the map.

STEP 4
When you hover over a country or use list view, you'll see an indexed search interest value between 1-100, just as you do with other Trends data.The location with the highest normalized interest in the topic valued at 100, while all other values are indexed proportionally.







Top and rising searches

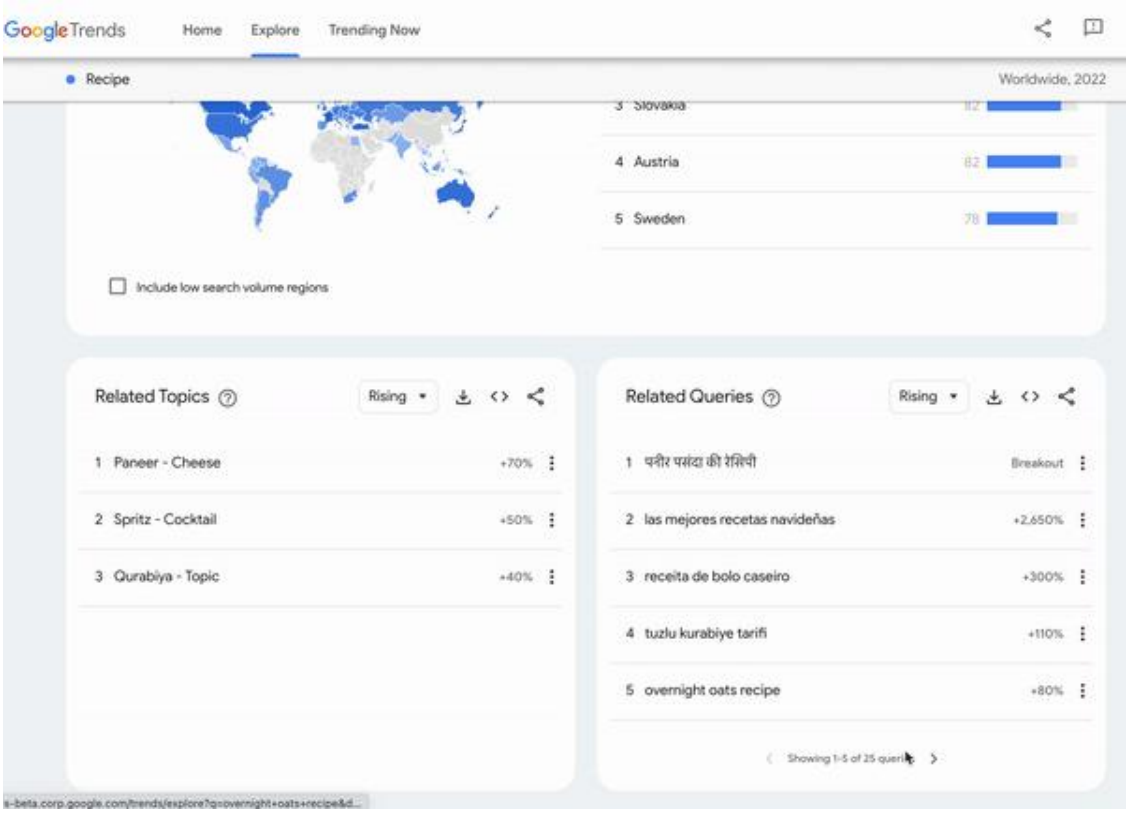
Top and rising search topics and queries make up the final key element of the Trends Explore tool. In this section, we'll learn how to read and understand this data.

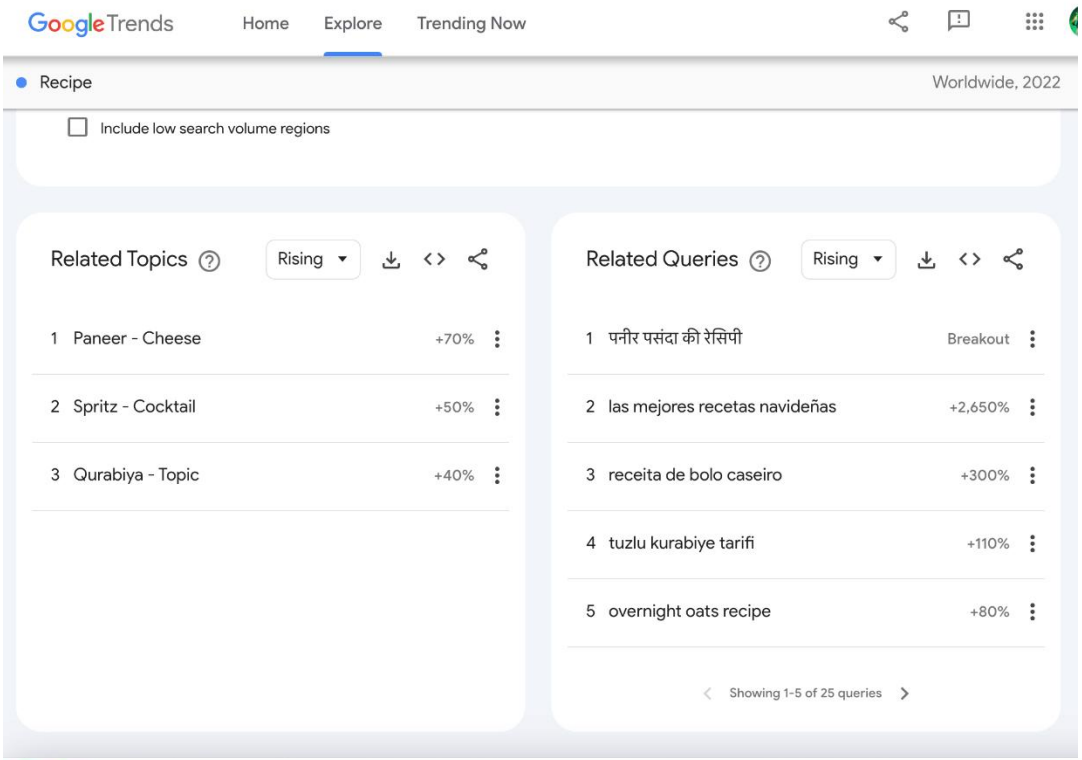
STEP 1
At the bottom of your Trends Explore page, you'll see tables of Related topics and Related queries pertaining to the term or topic you entered. To scroll through the results, click the arrows below the table.

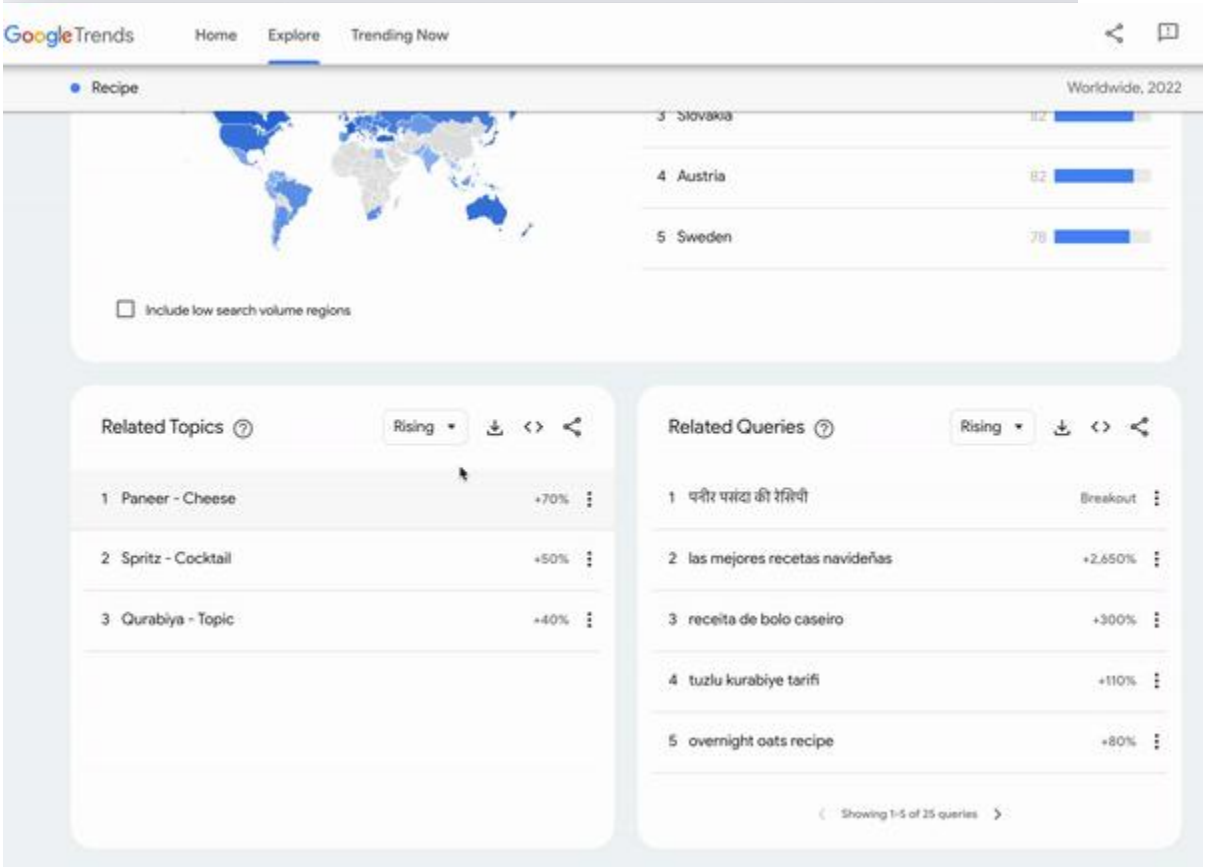
STEP 2
Trends Explore will automatically display "rising" data for both topics and queries. These have seen the largest increase in search interest over the selected time period. You can see the percent increase in search interest next to each topic or search term."Breakout" searches have increased by over 5000% percent in the given time frame. Usually these are new, or had very low search volume before.

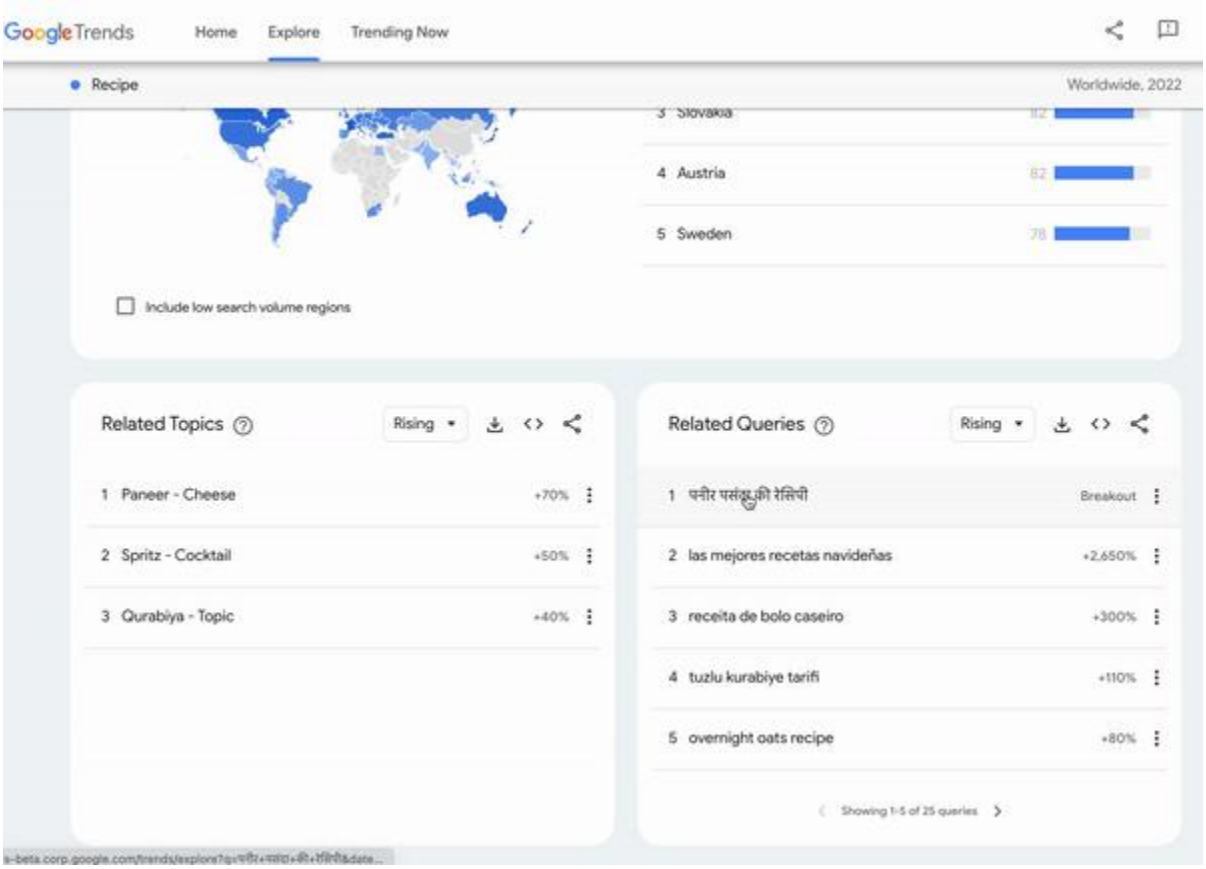
STEP 3
Click on the top right of the widget where it says Rising and switch to Top. These are the queries or topics with the highest search volume in the selected time frame and location. These are indexed and normalized values from 1-100.

STEP 4
Click on a top or rising topic or query to explore it in more detail.




Google Trends
Home
Explore
Trending Now
Recipe
Worldwide, 2022
Include low search volume regions
Related Topics
Rising
1 Paneer - Cheese +70%
2 Spritz - Cocktail +50%
3 Qurabiya - Topic +40%
Related Queries
Rising
1 पनीर पसंदा की रेसिपी Breakout
2 las mejores recetas navideñas +2,650%
3 receita de bolo caseiro +300%
4 tuzlu kurabiye tarifi +110%
5 overnight oats recipe +80%
Showing 1-5 of 25 queries

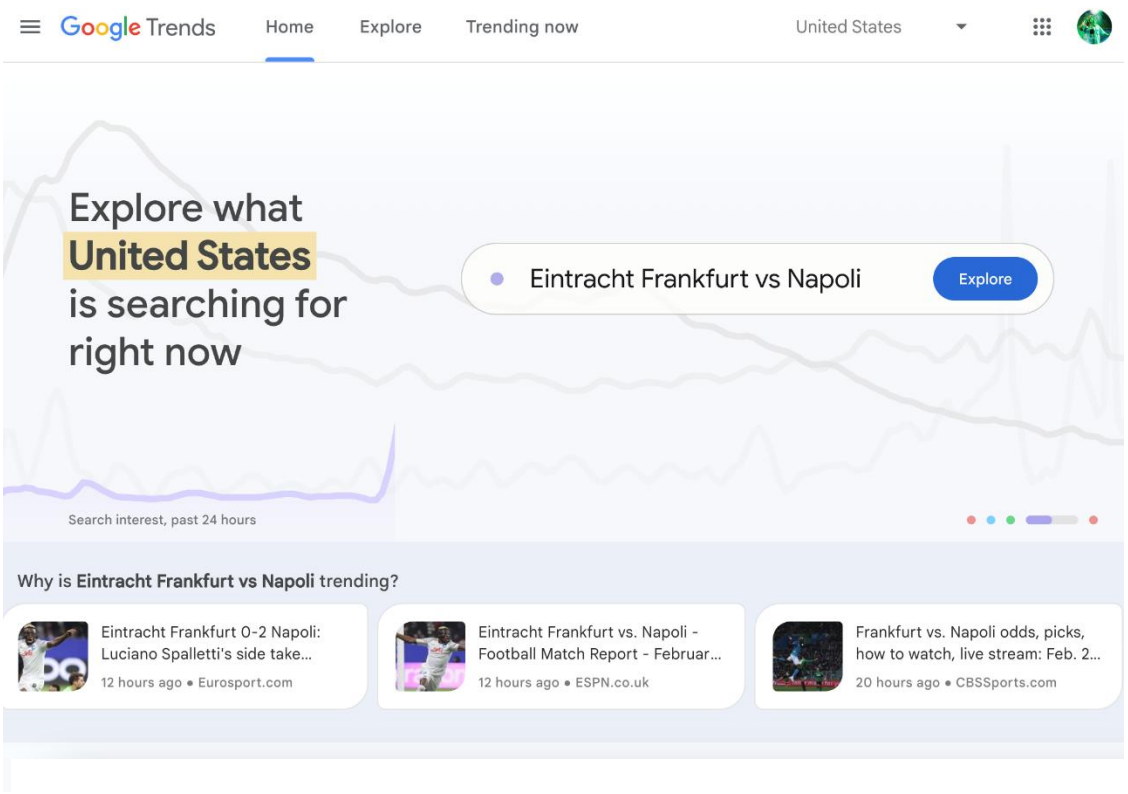

Google Trends
Home
Explore
Trending Now
Recipe
Worldwide, 2022
4 Austria
5 Sweden
Include low search volume regions
Related Topics
Rising
1 Paneer - Cheese +70%
2 Spritz - Cocktail +50%
3 Qurabiya - Topic +40%
Related Queries
Rising
1 पनीर पसंदा की रेसिपी Breakout
2 las mejores recetas navideñas +2,650%
3 receita de bolo caseiro +300%
4 tuzlu kurabiye tarifi +110%
5 overnight oats recipe +80%
Showing 1-5 of 25 queries



Data excluded from search

Trends excludes certain data from your searches.

- Searches made by very few people: Trends only analyzes data for popular terms, so search terms with low volume appear as 0 for a given time period.
- Duplicate searches: Trends eliminates repeated searches from the same user over a short period of time for better overall accuracy.
- Special characters: Trends filters out queries with apostrophes and other special characters.



SimilarWeb

SimilarWeb Ltd. is an Israeli web analytics company specializing in web traffic and performance.
The company was founded in 2007 by Or Offer in Tel Aviv, Israel. By 2009, Similarweb won the first Israeli SeedCamp, attracting the attention of international media and investors. The company raised its Series A round of $1.1 million[6] with the investment being led by Yossi Vardi, and Docor International Management. SimilarSites, a browser extension to help users find sites similar to those they are visiting, was launched later that year.
Similarweb develops tools that enable the analysis of the traffic and behavior of users on websites and apps. The service and datasets are provided in a limited free edition, but the paid platform is addressed to SMBs and large companies which require access to accurate comprehensive data at larger scales for marketing, sales and Market research. The data is collected from a number of different sources that provide information about the internet usage of users, including various information partners, and anonymous data from users of the various dedicated browser addons that the company distributes.

PART B
(PART B: TO BE COMPLETED BY STUDENTS)

Roll. No. 36	Name: Sanskruti Kadam
Class: BE C COMPS	Batch: C2
Date of Experiment:	Date of Submission:
Grade:	

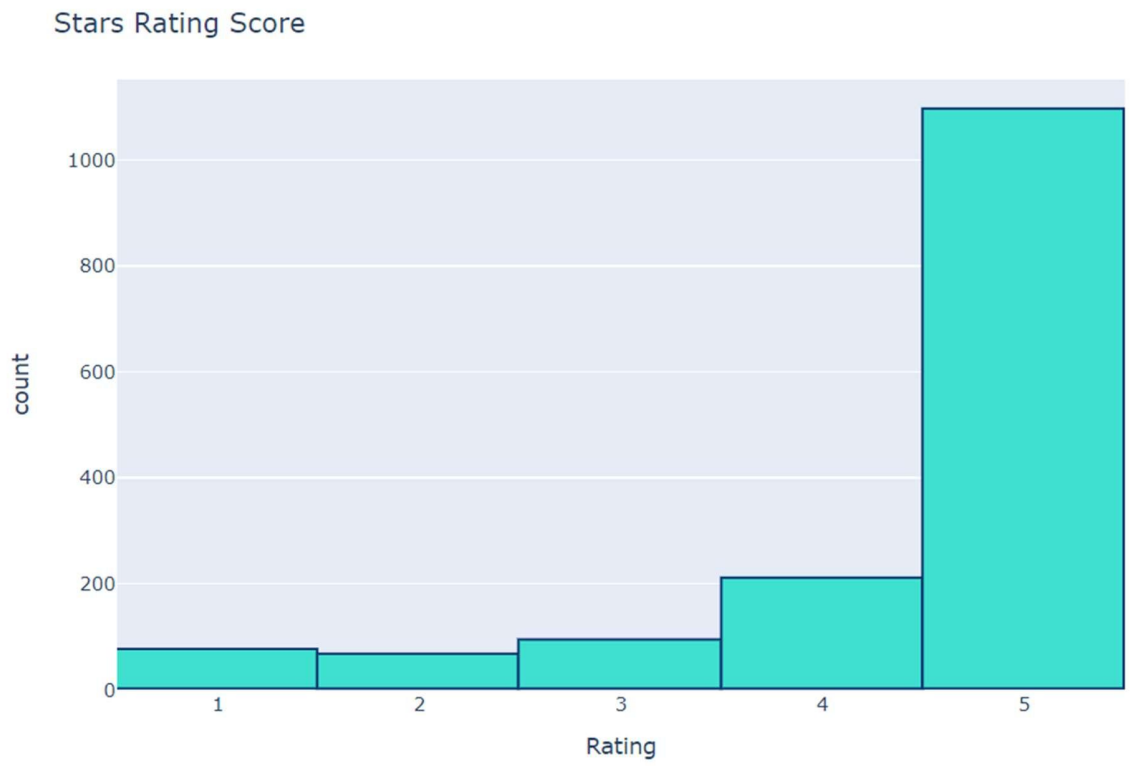
B.2 Input and Output:

```
import matplotlib.pyplot as plt
import seaborn as sns
color = sns.color_palette()
%matplotlib inline
import plotly.offline as py
py.init_notebook_mode(connected=True)
import  plotly.graph_objs  as  go
import plotly.tools as tls
import plotly.express as px
import pandas as pd
df = pd.read_csv('../input/product-reviews-and-ratings-sentiment-analysis/
Reviews And Ratings.csv')
df.head(10)
```

Out[4]:

	Unnamed: 0	Rating	Lang	Type	Country	Date	Helpful	translated
0	72	5	it	Verified	Italy	11/1/2020	43.0	What to say? My daughter LOVES him and I with ...
1	85	5	it	Not Verified	Italy	9/10/2019	36.0	I decided to test this plush to my grandson, j...
2	107	5	it	Verified	Italy	8/12/2019	2.0	An unusual pet, this otter is beautiful! Cute ...
3	109	4	it	Verified	Italy	4/2/2021	1.0	Beautiful, soft and very relaxing. It comes wi...
4	113	4	it	Verified	Italy	21/02/2020	2.0	Plush tender, my 16-month-old loves it, sleeps...
5	115	5	it	Verified	Italy	12/11/2019	3.0	I love it, simply. I bought it after seeing it...
6	116	5	it	Verified	Italy	25/10/2019	5.0	Beautiful product, the only problem and that t...
7	125	5	it	Verified	Italy	7/8/2020	2.0	Perfect!! Relaxed very much my son of 4 months...
8	138	5	it	Verified	Italy	25/04/2020	NaN	The otter accompanies the dwarfs of my baby wi...
9	146	2	it	Verified	Italy	23/11/2020	NaN	Functional carillon, my 4 month old girl falls...

```
fig = px.histogram(df, x="Rating")
fig.update_traces(marker_color="turquoise",marker_line_color='rgb(8,48,107
)',
                  marker_line_width=1.5)
fig.update_layout(title_text='Stars Rating Score')
fig.show()
```




```
import nltk
from nltk.corpus import stopwords
from spacy.lang.en.stop_words import STOP_WORDS
from wordcloud import WordCloud
# Create stopword list:
STOP_WORDS.add('otter')
stopwords = set(list(STOP_WORDS) +list(stopwords.words()))
stopwords.update(["br", "href", 'https'])
stopwords.update(stopwords)
textt = " ".join(desc for desc in df.translated)
wordcloud = WordCloud(stopwords=stopwords).generate(textt)
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.savefig('wordcloud11.png')
```

```
plt.show()
```


```
df = df[df['Rating'] != 0]
#Creating Positive & Negative sentiments as +1 and -1 according to rating
df['sentiment'] = df['Rating'].apply(lambda rating : +1 if rating >= 4 else -1)

positive = df[df['sentiment'] == 1]
negative = df[df['sentiment'] == -1]

nltk.download('stopwords')
```

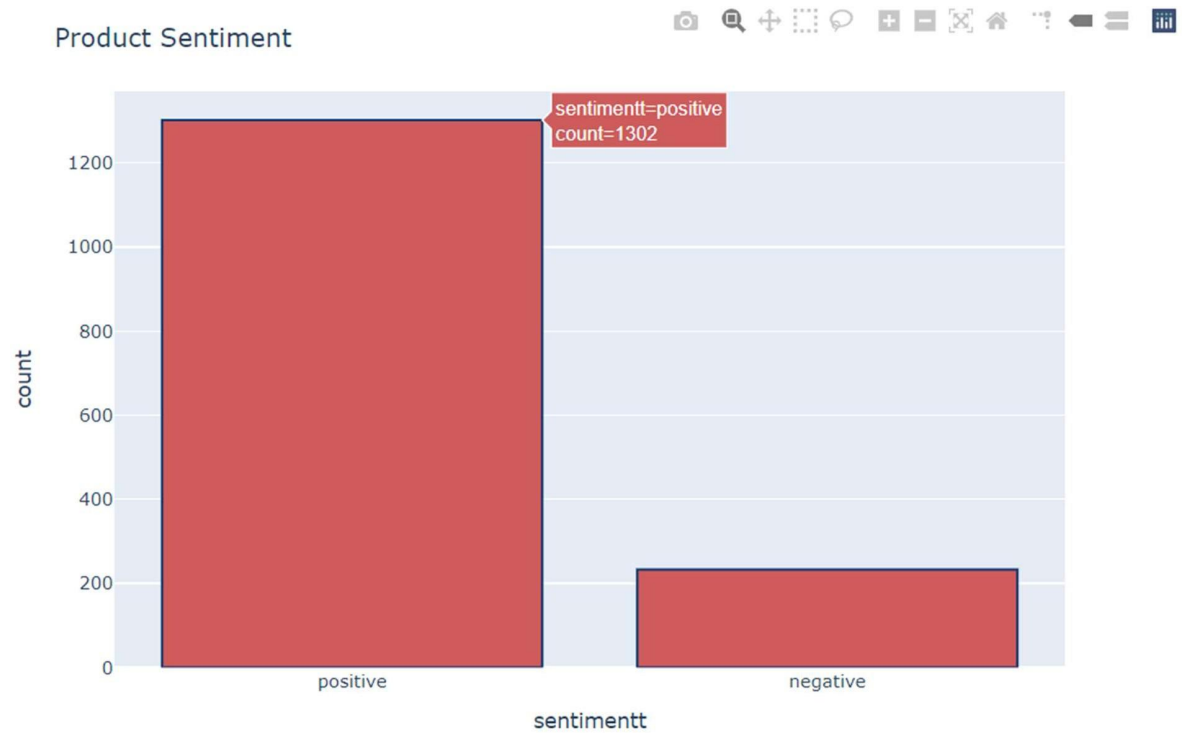
```
[nltk_data] Downloading package stopwords to /usr/share/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
```

```
stopwords.update(["br", "href","good","great", 'https'])
pos = " ".join(review for review in positive.translated)
wordcloud2 = WordCloud(stopwords=stopwords).generate(pos)
plt.imshow(wordcloud2, interpolation='bilinear')
plt.axis("off")
plt.show()
```


```
neg = " ".join(review for review in negative.translated)
wordcloud3 = WordCloud(stopwords=stopwords).generate(neg)
plt.imshow(wordcloud3, interpolation='bilinear')
plt.axis("off")
plt.savefig('wordcloud33.png')
plt.show()
```


```
df['sentimentt'] = df['sentiment'].replace({-1 : 'negative'})
df['sentimentt'] = df['sentimentt'].replace({1 : 'positive'})
```

```
fig = px.histogram(df, x="sentimentt")
fig.update_traces(marker_color="indianred",marker_line_color='rgb(8,48,107
)',
                  marker_line_width=1.5)
fig.update_layout(title_text='Product Sentiment')
fig.show()
```



B.3 Observations and learning:

The sentiment analysis revealed the overall sentiment expressed in the product reviews. Positive sentiment was predominant, indicating high customer satisfaction with the product. The experiment underscored the importance of data analytics in extracting meaningful information from large volumes of unstructured text data, enabling data-driven decision-making.

B.4 Conclusion:

The experiment involved performing text and sentiment analysis on Amazon product reviews to extract valuable insights.

B.5 Question of Curiosity

(To be answered by student based on the practical performed and learning/observations)

Q1. What is search engine analytics? What is the function of a search engine?

Search engine analytics refers to the process of analyzing and measuring data related to the performance and effectiveness of a website or web page in search engine results. It involves

tracking and interpreting various metrics to understand how users interact with a website through search engines like Google, Bing, Yahoo, etc.

The key functions of a search engine can be summarized as follows:

1. **Crawling**: Search engines use automated bots (crawlers or spiders) to browse and discover web pages across the internet. The crawlers follow links from one page to another, collecting data about each page they visit.
2. **Indexing**: After crawling, search engines organize and store the information they gather in a vast database called an index. The index includes details about the content and structure of web pages, making it easier to retrieve relevant results for user queries.
3. **Ranking**: When a user enters a search query, the search engine retrieves relevant pages from its index and ranks them based on various factors such as relevance, quality, authority, and user experience. Pages with higher rankings appear higher in the search results.
4. **Retrieval**: Finally, the search engine displays a list of search results (SERPs) that best match the user's query. Each result typically includes a title, URL, and snippet (short description) extracted from the indexed page.

Q2. Differentiate between local and global search engines?

Local Search Engines:

1. **Scope of Search Results**:
    - Local search engines prioritize providing search results that are relevant to a specific geographic location or region.
    - They focus on displaying businesses, services, and information that are relevant to users within a defined local area.
2. **Geographic Targeting**:
    - Local search engines target users based on their geographic location, typically identified through IP addresses or user-provided location data.
    - Search results are customized to display local businesses, addresses, contact information, and localized content.
3. **Use Case**:
    - Local search engines are commonly used for finding nearby businesses, restaurants, services, events, and local news.
    - Users often use local search engines to seek information specific to their city, neighborhood, or region.
4. **Examples**:
    - Google Maps: Provides local business listings, directions, and reviews based on user location.
    - Yelp: A platform focused on local business reviews and recommendations.
    - Bing Local: Bing's feature for discovering local businesses and services.

Global Search Engines:

1. **Scope of Search Results**:
    - Global search engines aim to provide search results that are relevant to users worldwide, irrespective of their geographic location.
    - They prioritize displaying information, websites, and content from across the globe without focusing on specific localities.
2. **Geographic Reach**:
    - Global search engines serve a broad and diverse audience, catering to users from various countries and regions.
    - They do not prioritize displaying location-specific results unless users specify their search intent (e.g., using location-specific keywords).
3. **Use Case**:
    - Global search engines are used for accessing general information, news, research, and global trends.
    - Users often use global search engines for academic research, international news, and exploring topics of global interest.
4. **Examples**:
    - Google Search: The most widely used global search engine providing access to information from across the internet.
    - Bing: Microsoft's search engine offering global search capabilities similar to Google.
    - Yahoo Search: Another global search engine providing access to a wide range of content.

Q3. What is the purpose of search engine trend analysis?

Here are the key purposes and benefits of search engine trend analysis:

1. **Identifying Popular Topics and Keywords**: Trend analysis helps in identifying which topics and keywords are currently trending or gaining popularity in search queries. This insight can inform content creation strategies, allowing businesses to create timely and relevant content that aligns with user interests.
2. **Forecasting Future Trends**: By analyzing historical search trends, businesses can forecast future trends and anticipate shifts in consumer behavior. This enables proactive planning and adaptation of marketing strategies to capitalize on emerging opportunities.
3. **Informing Content Strategy**: Search trend analysis guides content strategy by revealing the type of content users are actively searching for. This information can be used to develop content that resonates with the target audience and attracts organic search traffic.
4. **Optimizing SEO and PPC Campaigns**: Understanding search trends helps in optimizing search engine optimization (SEO) and pay-per-click (PPC) campaigns. Businesses can

prioritize keywords that are trending and adjust bidding strategies based on fluctuating search demand.

5. **Market Research and Consumer Insights**: Search trend data provides valuable insights into consumer preferences, behaviors, and needs. Businesses can use this data for market research purposes, product development, and identifying new business opportunities.
6. **Monitoring Competitor Activity**: Trend analysis allows businesses to monitor competitor activity by analyzing their search performance and identifying areas of opportunity or potential threats.
7. **Seasonal and Event-Based Campaigns**: By analyzing seasonal search trends and events (e.g., holidays, festivals), businesses can tailor marketing campaigns and promotions to capitalize on peak demand periods.
8. **Tracking Brand Performance**: Search trend analysis helps in monitoring brand performance by tracking brand-related search queries and sentiment over time. This data can inform brand reputation management strategies.

Q4. Explain the two main categories of search engine analytics.

On-page analytics: On-page analytics focuses on analyzing user interactions and behavior within individual web pages. It involves tracking metrics such as page views, bounce rates, time spent on page, and conversion rates. On-page analytics provides insights into how users engage with specific content, the effectiveness of design elements, and areas for optimization to improve user experience and achieve conversion goals.

Off-page analytics: Off-page analytics examines factors external to the website itself that influence search engine visibility and performance. This includes backlink analysis, social media signals, brand mentions, and overall online reputation. Off-page analytics helps businesses understand their authority and relevance within the broader online ecosystem, identify opportunities for link building and partnership, and assess the impact of external factors on search engine rankings.