

# DATA VISUALIZATION AND SQL INTRO : KAGGLE.COM COURSE

CONTENT DERIVED FROM (AND MODIFIED) :KAGGLE.COM  
COURSES(APACHE 2.0 OPEN SOURCE LICENSE)

- We use python most of the either through the notebook or through scrips. So, what are data visualization and what's their functioning. Data Visualization are tools and techniques to visualize the data. It can be between ordinate and coordinate or b/w multiples variables.
- Typically , in python seaborn is used for data visualization and matplotlib is another common library used for plotting maps and visualization in python.
- For Data visualization ,seaborn is imported as sns.

# DIFFERENT PLOTS AVAILABLE IN THE FORM SNS.LINECHART()

- LINEPLOT()  
BAR CHART()  
SCATTERPLOT()  
HEATMAP()  
HISTOGRAM()  
SMOOTHED HISTOGRAM(DENSITY PLOT()).
- MANY MORE.

# BIG QUERY: STRUCTURED QUERY LANGUAGE

- `From.google.cloud.import Big Query.`
- Initial steps for query:
- Create a client object.
- To access the dataset, create a reference to dataset in the construction `client.dataset()`.
- Fetch the dataset with the construction created as argument and `client.get_dataset()` function.
- `List(client.list_tables(dataset))` is use to list the tables in the dataset.
- `Table.scheme` is used to see structure of table i.e. structure of table.

# SELECT FROM WHERE

- Dealing with dataset in SQL is done with the help of queries. SELECT, FROM AND WHERE ARE EXAMPLE OF QUERIES IN SQL.
- SELECT...FROM.  
Specify the column we want to take data using the word SELECT and then FROM is used to specify the table for the data set to get selected for querying in SQL.  
To select a column from a dataset where a specific condition on row value is to be met , WHERE is used such as WHERE COLOR = WHITE;

- GROUP BY AND HAVING.
- By grouping columns particular value in row we use GROUP BY and Having is used to specify a particular condition on groupby or to ignore groups that don't meet a particular condition.
- Count is a aggregate function and used to count the things ,similar function of aggregate form are SUM(), AVG(),MIN(),MAX().
- For using GROUP BY function, all the variables have been passed through either an aggregation function or a GROUP BY command.

- ORDER BY:
- A typical SQL query starts with query = ""

```
SELECT id,  
queries etc.  
""
```

ORDER BY is one of the important query of the SQL and its is used to sort the result after the all queries are done , it can be by sone row value such as id or desc / aesc ordering.



- **Aliasing:**

- AS is used for aliasing similar to python libraries name, it is used to aliasing the column name most of the times , after the SELECT ID as A2.