Low Level Design (LLD)

Entertainer Data Analysis

Bhakti Arvind Khavle
Prathmesh Narayan Khapare

Index

Abstract

- Normal life can be stressful, and people need to relax. Being entertained by others is a wonderful way to take some time out of life. It can reduce stress and make life's issues easier to face. The media and entertainment industry consists of film, television, radio and print. These segments include movies, TV shows, radio shows, news, music, newspapers, magazines, and books. Entertainment industry is a group of sub-industries devoted to entertainment. Entertainment industry is used to describe the mass media companies that control the distribution and manufacture of mass media entertainment.

Given Tasks

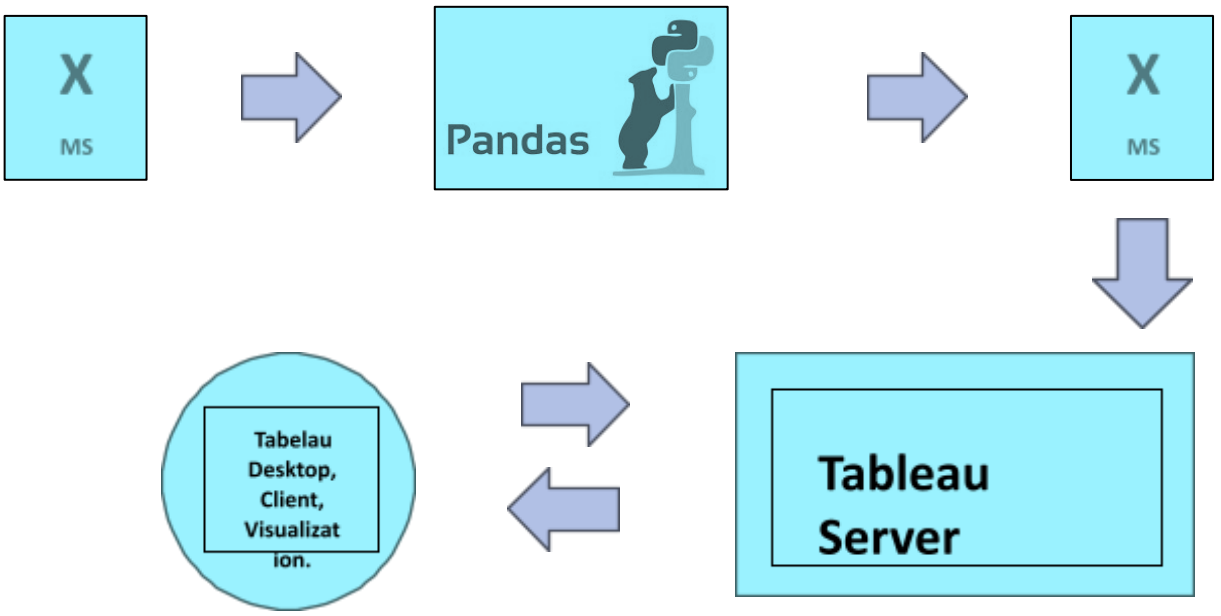
- **Task #1 -** In a word document write the process and data added to the current dataset. In addition, mention the theme on which you will be creating the dashboard.
    - **Task #2 -** You can add your data as per your convenience.
    - **Task #3** – Do the data preparation part.
    - **Task #4** – Build the dashboards
    - **Task #5** – Build a Storyline

Scope

- My main theme is the number of awards won by entertainers throughout their life.I have made two dashboards showing their award winning performance. Dashboards show the highest number of nominees, Oscar awards, Emmy awards, Grammy awards and other awards etc. so that each and every one can analyze which entertainer is the best according to their thought process.
- My main theme is to create dashboards on top entertainers with their highest number of awards and nominees and make visualizations on it.

Architecture

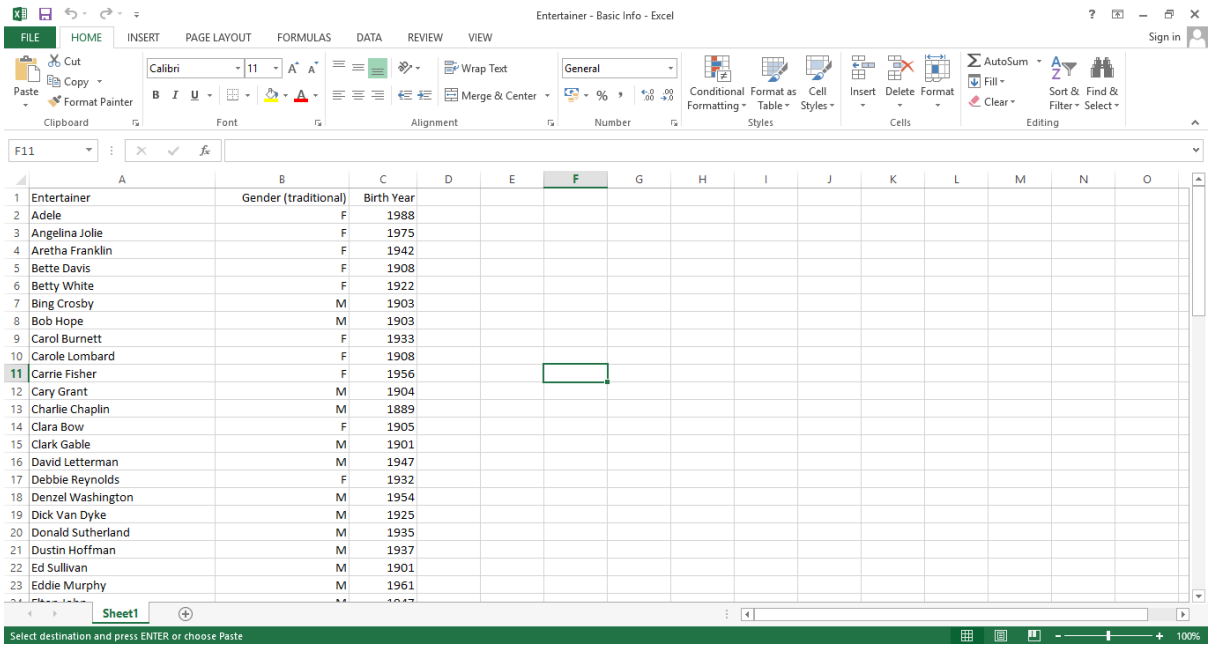
- The **architecture** of entire project is shown below:



- Our entire data source is our **Excel** file.
- After **Excel** I am pushing this dataset in **Pandas** (Python library) in pandas in manipulated the tables and **merged** them into one table using the .**concat** function. Then in converted that table into excel file to push into **Tableau** server to make a **dashboard**.
- This excel file is connected to the **Tableau** server. From the server, data can be shown and **accessed**
- **Tableau** server has various architectural components regarding how to solve the query.
- The functionalities show the result according to a query entered by the end user or client.
- Screen of **Tableau** desktop, client and various charts and **dashboard** (screen) of Tableau are present at client side.
- Client entered the query to show the **graph**. After selecting the data in the form of rows and columns it will go inside the tableau server. In the tableau server, it understands the query and generates the best recommended charts based on selected data and returns it into the tableau screen.
- Based on recommended charts, clients can make the visual aspect the same.
- If a client is not **satisfied** with the result, he/she has to select data accordingly otherwise make required changes to show the expected result.

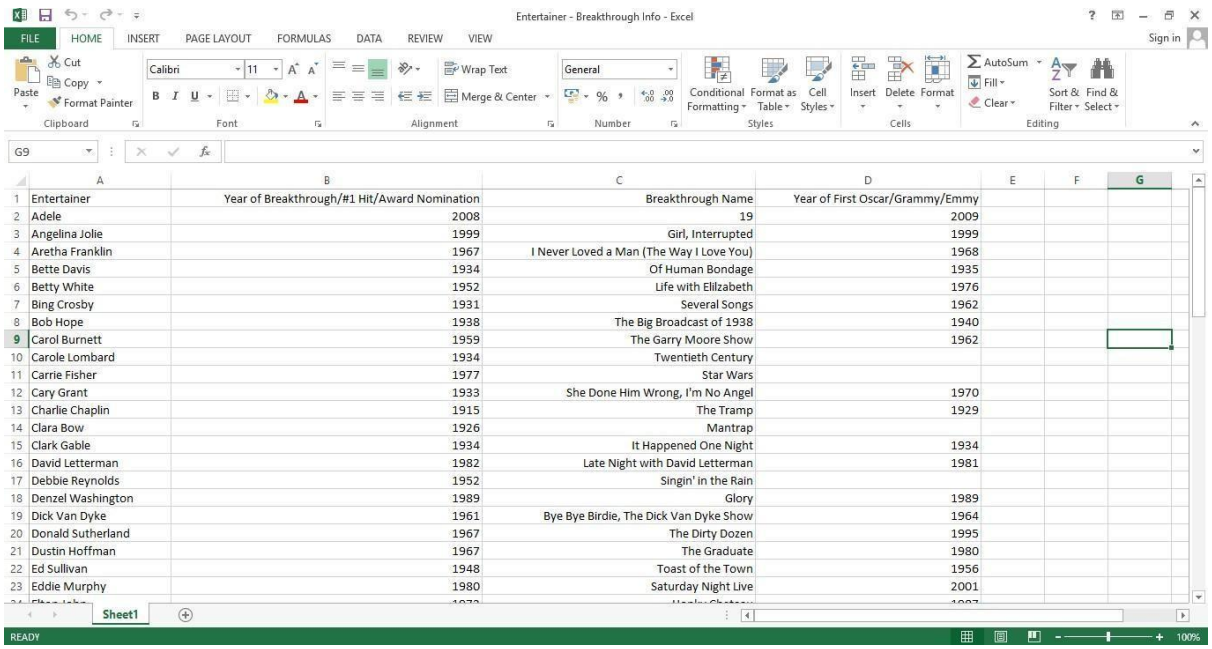
Data Description

- Data was given into three parts in excel file which are Entertainer – Basic Info, Entertainer – Breakthrough Info and Entertainer - Last work Info.
- Basic Info includes name, gender and birth year of the entertainer.
- Breakthrough Info includes year of breakthrough/#1 hit/award nomination, breakthrough name, year of first Oscar/grammy/emmy along with name of entertainers.
- Last work Info includes last major work (arguable), year of death (if they are) along with names of entertainers.
- Glimpse of Entertainer – **Basic Info**:



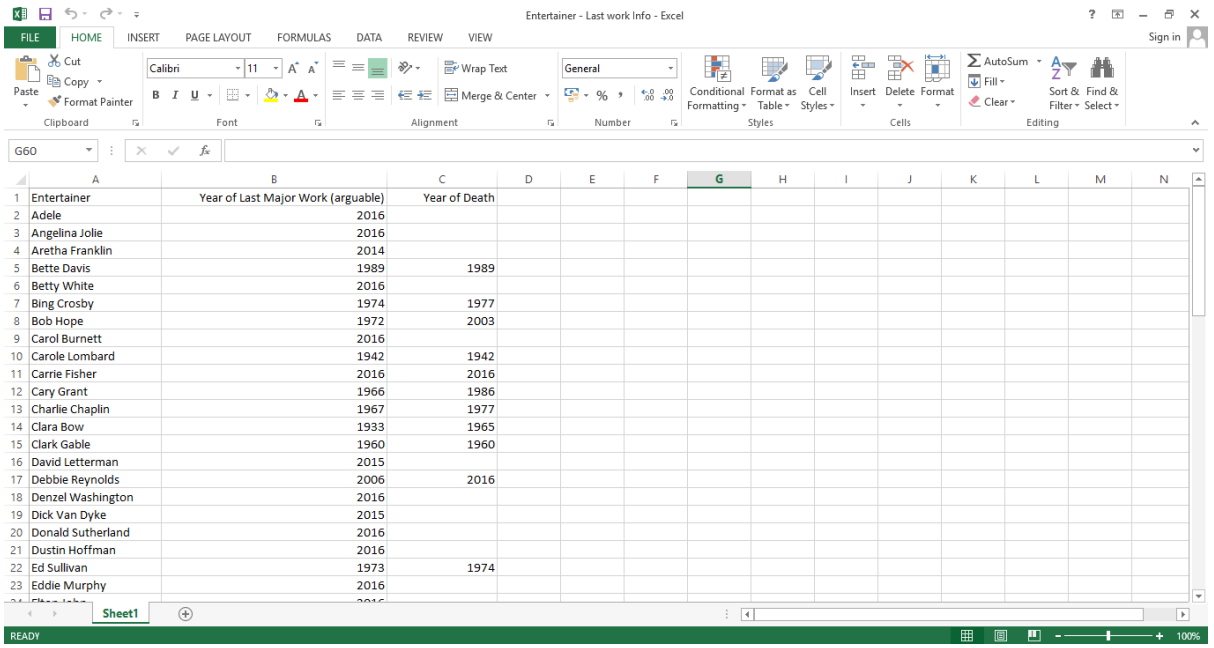
	Entertainer	Gender (traditional)	Birth Year
2	Adele	F	1988
3	Angelina Jolie	F	1975
4	Aretha Franklin	F	1942
5	Bette Davis	F	1908
6	Betty White	F	1922
7	Bing Crosby	M	1903
8	Bob Hope	M	1903
9	Carol Burnett	F	1933
10	Carole Lombard	F	1908
11	Carrie Fisher	F	1956
12	Cary Grant	M	1904
13	Charlie Chaplin	M	1889
14	Clara Bow	F	1905
15	Clark Gable	M	1901
16	David Letterman	M	1947
17	Debbie Reynolds	F	1932
18	Denzel Washington	M	1954
19	Dick Van Dyke	M	1925
20	Donald Sutherland	M	1935
21	Dustin Hoffman	M	1937
22	Ed Sullivan	M	1901
23	Eddie Murphy	M	1961

- Glimpse of Entertainer – **Breakthrough Info**:



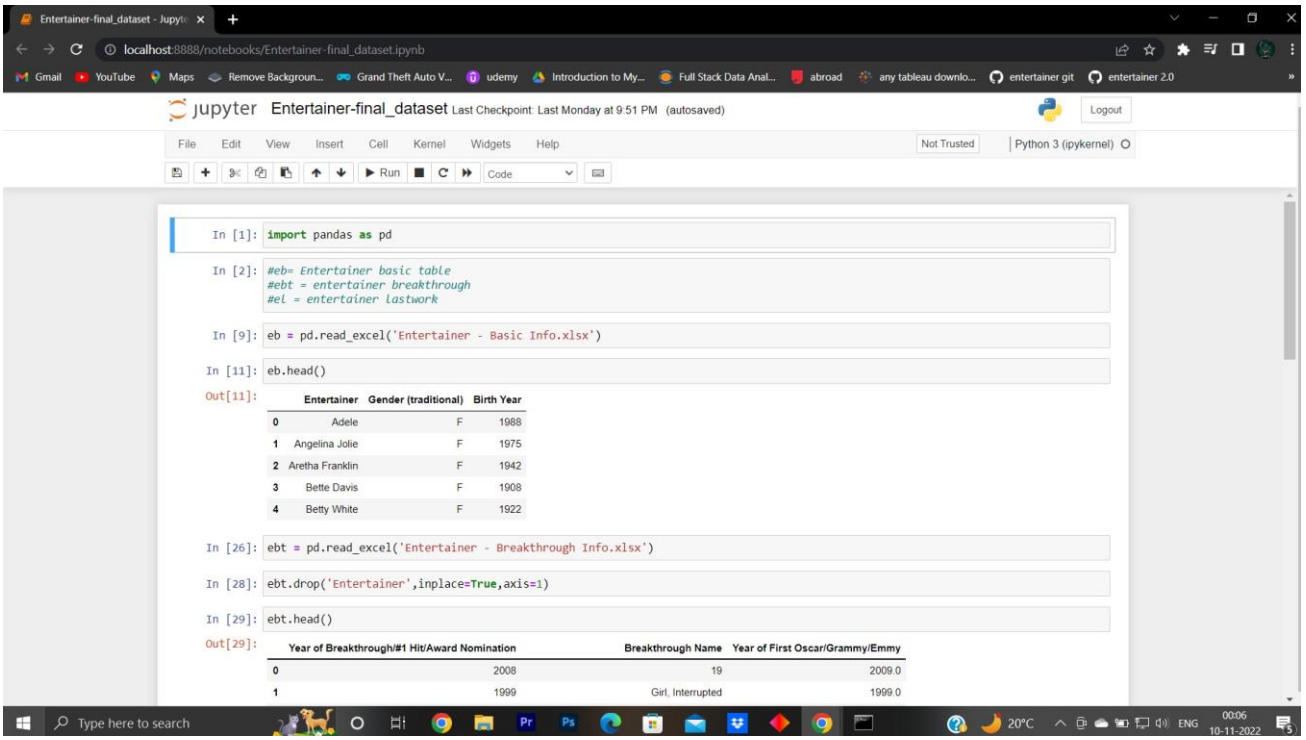
	Entertainer	Year of Breakthrough/#1 Hit/Award Nomination	Breakthrough Name	Year of First Oscar/Grammy/Emmy
2	Adele	2008	19	2009
3	Angelina Jolie	1999	Girl, Interrupted	1999
4	Aretha Franklin	1967	I Never Loved a Man (The Way I Love You)	1968
5	Bette Davis	1934	Of Human Bondage	1935
6	Betty White	1952	Life with Elilzabeth	1976
7	Bing Crosby	1931	Several Songs	1962
8	Bob Hope	1938	The Big Broadcast of 1938	1940
9	Carol Burnett	1959	The Garry Moore Show	1962
10	Carole Lombard	1934	Twentieth Century	
11	Carrie Fisher	1977	Star Wars	
12	Cary Grant	1933	She Done Him Wrong, I'm No Angel	1970
13	Charlie Chaplin	1915	The Tramp	1929
14	Clara Bow	1926	Mantrap	
15	Clark Gable	1934	It Happened One Night	1934
16	David Letterman	1982	Late Night with David Letterman	1981
17	Debbie Reynolds	1952	Singin' in the Rain	
18	Denzel Washington	1989	Glory	1989
19	Dick Van Dyke	1961	Bye Bye Birdie, The Dick Van Dyke Show	1964
20	Donald Sutherland	1967	The Dirty Dozen	1995
21	Dustin Hoffman	1967	The Graduate	1980
22	Ed Sullivan	1948	Toast of the Town	1956
23	Eddie Murphy	1980	Saturday Night Live	2001

➢ Glimpse of Entertainer – **Last work Info**;



	A	B	C
1	Entertainer	Year of Last Major Work (arguable)	Year of Death
2	Adele	2016	
3	Angelina Jolie	2016	
4	Aretha Franklin	2014	
5	Bette Davis	1989	1989
6	Betty White	2016	
7	Bing Crosby	1974	1977
8	Bob Hope	1972	2003
9	Carol Burnett	2016	
10	Carole Lombard	1942	1942
11	Carrie Fisher	2016	2016
12	Cary Grant	1966	1986
13	Charlie Chaplin	1967	1977
14	Clara Bow	1933	1965
15	Clark Gable	1960	1960
16	David Letterman	2015	
17	Debbie Reynolds	2006	2016
18	Denzel Washington	2016	
19	Dick Van Dyke	2015	
20	Donald Sutherland	2016	
21	Dustin Hoffman	2016	
22	Ed Sullivan	1973	1974
23	Eddie Murphy	2016	

➢ After this, I combined this data into one Table using Pandas as this screenshot describes how I merged the tables using .**CONCAT** function .As this screenshot shows I worked in **Jupyter Notebook** as it is easy and perfect for **Data Analysis** and **Data Science .** And named this file as **Entertainer-final_dataset**



```
In [1]: import pandas as pd

In [2]: #eb= Entertainer basic table
        #ebt = entertainer breakthrough
        #el = entertainer lastwork

In [9]: eb = pd.read_excel('Entertainer - Basic Info.xlsx')

In [11]: eb.head()
```

Out[11]:

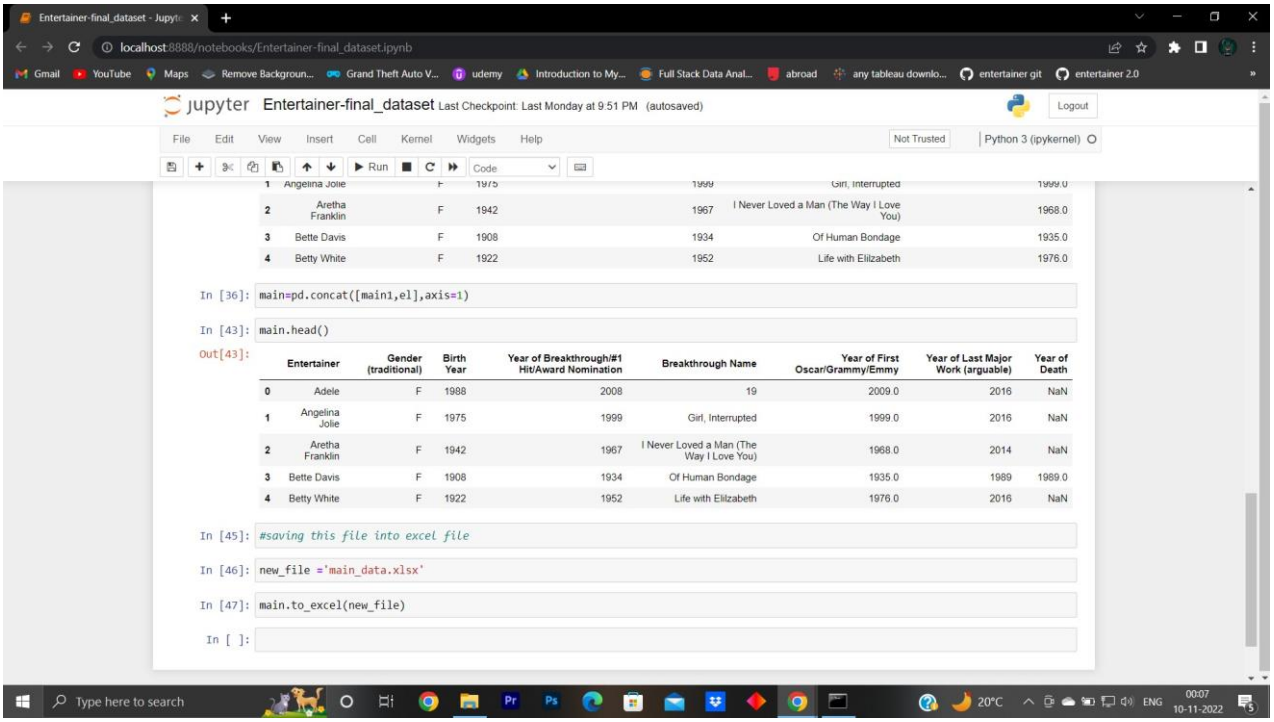
	Entertainer	Gender (traditional)	Birth Year
0	Adele	F	1988
1	Angelina Jolie	F	1975
2	Aretha Franklin	F	1942
3	Bette Davis	F	1908
4	Betty White	F	1922

```
In [26]: ebt = pd.read_excel('Entertainer - Breakthrough Info.xlsx')

In [28]: ebt.drop('Entertainer',inplace=True,axis=1)

In [29]: ebt.head()
```

Out[29]:

	Year of Breakthrough/#1 Hit/Award Nomination	Breakthrough Name	Year of First Oscar/Grammy/Emmy
0	2008	19	2009.0
1	1999	Girl, Interrupted	1999.0

- Here the image shows how I saved the table in **excel** format as it is easy to apply in the **tableau** server and build the **dashboard** as per our convenience. As named **main_data.xlsx**

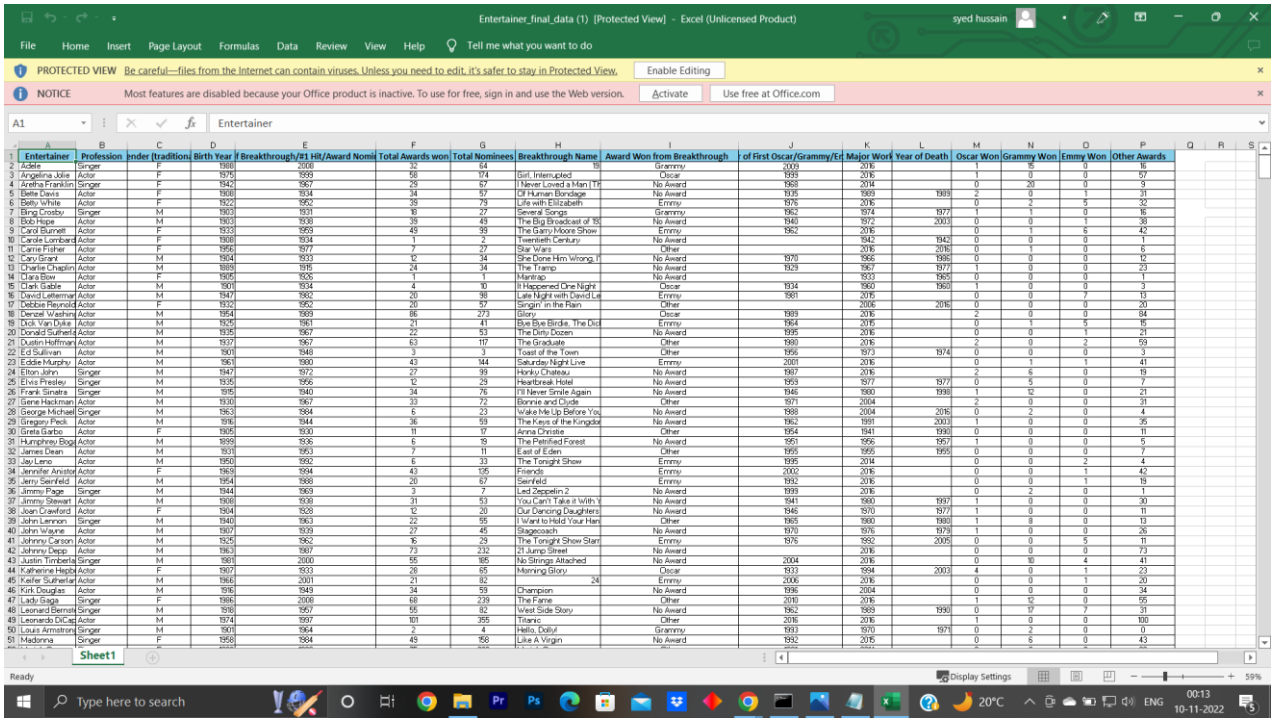


```
In [36]: main=pd.concat([main1,el],axis=1)

In [43]: main.head()

Out[43]:
```

	Entertainer	Gender (traditional)	Birth Year	Year of Breakthrough/#1 Hit/Award Nomination	Breakthrough Name	Year of First Oscar/Grammy/Emmy	Year of Last Major Work (arguable)	Year of Death
0	Adele	F	1988	2008	19	2009.0	2016	NaN
1	Angelina Jolie	F	1975	1999	Girl, Interrupted	1999.0	2016	NaN
2	Aretha Franklin	F	1942	1967	I Never Loved a Man (The Way I Love You)	1968.0	2014	NaN
3	Bette Davis	F	1908	1934	Of Human Bondage	1935.0	1989	1989.0
4	Betty White	F	1922	1952	Life with Elilzabeth	1976.0	2016	NaN

```
In [45]: #saving this file into excel file

In [46]: new_file ='main_data.xlsx'

In [47]: main.to_excel(new_file)

In [ ]:
```

- After this, I combined this data into one sheet named **Entertainer-final-data,** but this data was not sufficient to do analysis and make a dashboard on it. For this, I have added some external information like awards they won, nominees etc. from **IMDb's** official website (IMDb).



- *Then I pushed the data into the Tableau server as you will see in the last page.*

- **Entertainer**: Name of the entertainer.

- **Gender (traditional)**: Gender of that entertainer

- **Birth Year**: Birth year of that entertainer

- **Year of Breakthrough/#1 hit/Award Nominatio**n: Here, breakthrough means super hit or career changing performance. Column shows year of breakthrough

- **Breakthrough Name:** Name of breakthrough. It can be either a musical album or TV show or movie.

- **Year of first Oscar/Grammy/Emmy:** Year of first mega award they won.

- **Year of Last Major Work (arguable):** Last major show or movie or album. You can also say last appearance.
- **Year of Death:** Entertainer's year of death, if they die.

- **Award won from Breakthrough:** Any award/s from breakthrough. I only wrote about the mega awards they won in this column. If they had other awards for breakthroughs, I wrote "other", if they didn't, I wrote "No Award". If they have a mega award, I wrote that award name.

- **Total Awards Won:** Total awards that entertainer won throughout.

- **Total Nominees:** Total nominees for which they have chosen for award.

- **Profession:** Category of entertainer either singer or actor. Pop stars and dancers are included in singers and TV hosts and TV actors are included in actors.

- **Oscar won:** Number of total Oscar awards they won.

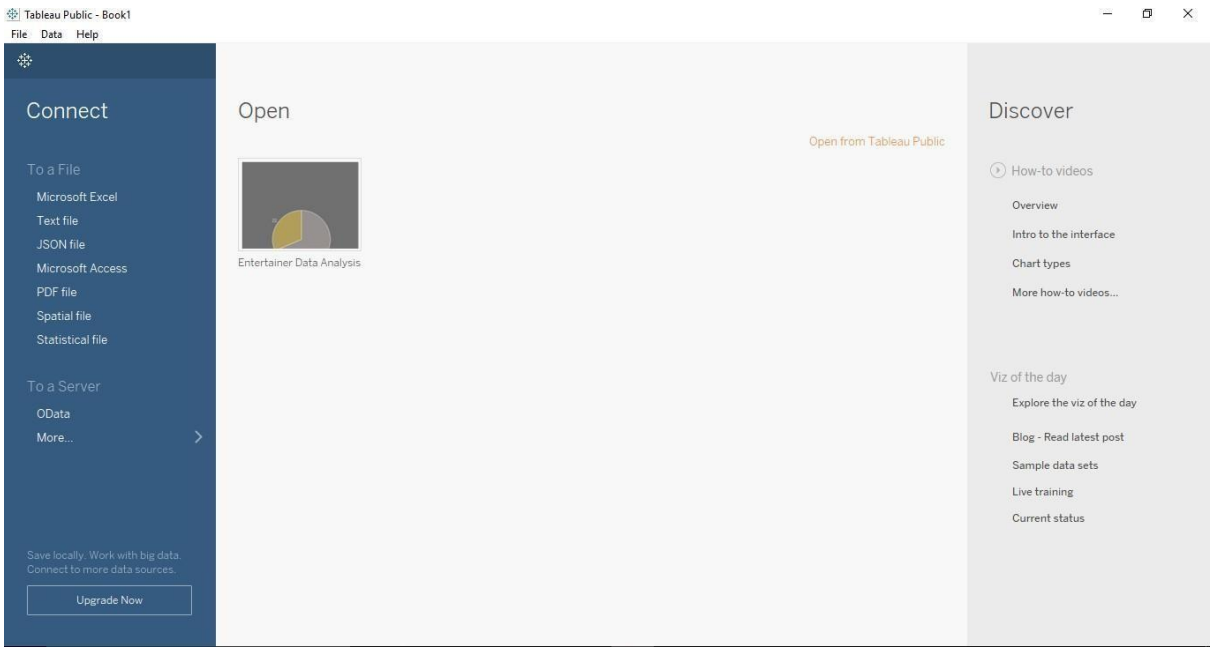
- **Grammy won:** Number of total Grammy awards they won.

- **Emmy won:** Number of total Emmy awards they won.

- **Other Awards:** Number of other awards they won apart from Oscar, Grammy and Emmy

Connect Data with Tableau & Deployment

- First of all, open Tableau Public on your desktop. At the first screen, it will ask you to connect your files from various sources like MS Excel, SQL Server, Tableau Serveretc.
- First screen of Tableau looks like:



- Make sure the internet connection is connected well while working with tableau, otherwise it will show the error.
- After completion of work, you can simply press ctrl + s or save it from the file menu. It will let you to tableau public's website and ask you for signing in. After sign-in, your work will be saved on tableau's website. There, all can see the work.
- In **Tableau** we can make professional dashboards as you can in the dashboards provided to you on Tableau public server