

EDA Automation (First Finding AI based Insights)

```
pip install gradio
```

```
Requirement already satisfied: gradio in c:\anaconda3\lib\site-packages (5.23.3)
```

```
Requirement already satisfied: aiofiles<24.0,>=22.0 in c:\anaconda3\lib\site-packages (from gradio) (23.2.1)
```

```
Requirement already satisfied: anyio<5.0,>=3.0 in c:\anaconda3\lib\site-packages (from gradio) (4.2.0)
```

```
Requirement already satisfied: fastapi<1.0,>=0.115.2 in c:\anaconda3\lib\site-packages (from gradio) (0.115.12)
```

```
Requirement already satisfied: ffmpy in c:\anaconda3\lib\site-packages (from gradio) (0.5.0)
```

```
Requirement already satisfied: gradio-client==1.8.0 in c:\anaconda3\lib\site-packages (from gradio) (1.8.0)
```

```
Requirement already satisfied: groovy~=0.1 in c:\anaconda3\lib\site-packages (from gradio) (0.1.2)
```

```
Collecting httpx>=0.24.1 (from gradio)
```

```
Using cached httpx-0.28.1-py3-none-any.whl.metadata (7.1 kB)
```

```
Requirement already satisfied: huggingface-hub>=0.28.1 in c:\anaconda3\lib\site-packages (from gradio) (0.30.1)
```

```
Requirement already satisfied: jinja2<4.0 in c:\anaconda3\lib\site-packages (from gradio) (3.1.4)
```

```
Requirement already satisfied: markupsafe<4.0,>=2.0 in c:\anaconda3\lib\site-packages (from gradio) (2.1.3)
```

```
Requirement already satisfied: numpy<3.0,>=1.0 in c:\anaconda3\lib\site-packages (from gradio) (1.26.4)
```

```
Requirement already satisfied: orjson~=3.0 in c:\anaconda3\lib\site-packages (from gradio) (3.10.16)
```

```
Requirement already satisfied: packaging in c:\anaconda3\lib\site-packages (from gradio) (24.1)
```

```
Requirement already satisfied: pandas<3.0,>=1.0 in c:\anaconda3\lib\site-packages (from gradio) (2.2.2)
```

```
Requirement already satisfied: pillow<12.0,>=8.0 in c:\anaconda3\lib\site-packages (from gradio) (10.4.0)
```

```
Requirement already satisfied: pydantic<2.12,>=2.0 in c:\anaconda3\lib\site-packages (from gradio) (2.11.2)
```

```
Requirement already satisfied: pydub in c:\anaconda3\lib\site-packages (from gradio) (0.25.1)
```

```
Requirement already satisfied: python-multipart>=0.0.18 in c:\anaconda3\lib\site-packages (from gradio) (0.0.20)
```

```
Requirement already satisfied: pyyaml<7.0,>=5.0 in c:\anaconda3\lib\site-packages (from gradio) (6.0.1)
```

```
Requirement already satisfied: ruff>=0.9.3 in c:\anaconda3\lib\site-packages (from gradio) (0.11.4)
```

```
Requirement already satisfied: safehttpx<0.2.0,>=0.1.6 in c:\anaconda3\lib\site-packages (from gradio) (0.1.6)
```

```
Requirement already satisfied: semantic-version~=2.0 in c:\anaconda3\
```

lib\site-packages (from gradio) (2.10.0)
Requirement already satisfied: starlette<1.0,>=0.40.0 in c:\anaconda3\lib\site-packages (from gradio) (0.46.1)
Requirement already satisfied: tomlkit<0.14.0,>=0.12.0 in c:\anaconda3\lib\site-packages (from gradio) (0.13.2)
Requirement already satisfied: typer<1.0,>=0.12 in c:\anaconda3\lib\site-packages (from gradio) (0.15.2)
Requirement already satisfied: typing-extensions~=4.0 in c:\anaconda3\lib\site-packages (from gradio) (4.13.1)
Requirement already satisfied: uvicorn>=0.14.0 in c:\anaconda3\lib\site-packages (from gradio) (0.34.0)
Requirement already satisfied: fsspec in c:\anaconda3\lib\site-packages (from gradio-client==1.8.0->gradio) (2024.6.1)
Requirement already satisfied: websockets<16.0,>=10.0 in c:\anaconda3\lib\site-packages (from gradio-client==1.8.0->gradio) (10.4)
Requirement already satisfied: idna>=2.8 in c:\anaconda3\lib\site-packages (from anyio<5.0,>=3.0->gradio) (2.10)
Requirement already satisfied: sniffio>=1.1 in c:\anaconda3\lib\site-packages (from anyio<5.0,>=3.0->gradio) (1.3.0)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\anaconda3\lib\site-packages (from pandas<3.0,>=1.0->gradio) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in c:\anaconda3\lib\site-packages (from pandas<3.0,>=1.0->gradio) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in c:\anaconda3\lib\site-packages (from pandas<3.0,>=1.0->gradio) (2023.3)
Requirement already satisfied: annotated-types>=0.6.0 in c:\anaconda3\lib\site-packages (from pydantic<2.12,>=2.0->gradio) (0.6.0)
Requirement already satisfied: pydantic-core==2.33.1 in c:\anaconda3\lib\site-packages (from pydantic<2.12,>=2.0->gradio) (2.33.1)
Requirement already satisfied: typing-inspection>=0.4.0 in c:\anaconda3\lib\site-packages (from pydantic<2.12,>=2.0->gradio) (0.4.0)
Requirement already satisfied: click>=8.0.0 in c:\anaconda3\lib\site-packages (from typer<1.0,>=0.12->gradio) (8.2.1)
Requirement already satisfied: shellingham>=1.3.0 in c:\anaconda3\lib\site-packages (from typer<1.0,>=0.12->gradio) (1.5.0)
Requirement already satisfied: rich>=10.11.0 in c:\anaconda3\lib\site-packages (from typer<1.0,>=0.12->gradio) (13.7.1)
Requirement already satisfied: colorama in c:\anaconda3\lib\site-packages (from click>=8.0.0->typer<1.0,>=0.12->gradio) (0.4.6)
Requirement already satisfied: certifi in c:\anaconda3\lib\site-packages (from httpx>=0.24.1->gradio) (2025.1.31)
Collecting httpcore==1.* (from httpx>=0.24.1->gradio)
Using cached httpcore-1.0.9-py3-none-any.whl.metadata (21 kB)
Collecting h11>=0.16 (from httpcore==1.*->httpx>=0.24.1->gradio)
Using cached h11-0.16.0-py3-none-any.whl.metadata (8.3 kB)
Requirement already satisfied: filelock in c:\anaconda3\lib\site-packages (from huggingface-hub>=0.28.1->gradio) (3.13.1)
Requirement already satisfied: requests in c:\anaconda3\lib\site-packages (from huggingface-hub>=0.28.1->gradio) (2.32.3)

```

Requirement already satisfied: tqdm>=4.42.1 in c:\anaconda3\lib\site-
packages (from huggingface-hub>=0.28.1->gradio) (4.67.1)
Requirement already satisfied: six>=1.5 in c:\anaconda3\lib\site-
packages (from python-dateutil>=2.8.2->pandas<3.0,>=1.0->gradio)
(1.16.0)
Requirement already satisfied: markdown-it-py>=2.2.0 in c:\anaconda3\
lib\site-packages (from rich>=10.11.0->typer<1.0,>=0.12->gradio)
(2.2.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in c:\
anaconda3\lib\site-packages (from rich>=10.11.0->typer<1.0,>=0.12-
>gradio) (2.15.1)
Requirement already satisfied: mdurl~=0.1 in c:\anaconda3\lib\site-
packages (from markdown-it-py>=2.2.0->rich>=10.11.0->typer<1.0,>=0.12-
>gradio) (0.1.0)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\
anaconda3\lib\site-packages (from requests->huggingface-hub>=0.28.1-
>gradio) (3.3.2)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\anaconda3\lib\
site-packages (from requests->huggingface-hub>=0.28.1->gradio)
(1.26.20)
Using cached httpx-0.28.1-py3-none-any.whl (73 kB)
Using cached httpcore-1.0.9-py3-none-any.whl (78 kB)
Using cached h11-0.16.0-py3-none-any.whl (37 kB)
Installing collected packages: h11, httpcore, httpx

```

```

    Attempting uninstall: h11

```

```

      Found existing installation: h11 0.12.0

```

```

    Uninstalling h11-0.12.0:

```

```

      Successfully uninstalled h11-0.12.0

```

```

----- 0/3 [h11]
Attempting uninstall: httpcore
----- 0/3 [h11]
    Found existing installation: httpcore 0.13.3
----- 0/3 [h11]
    Uninstalling httpcore-0.13.3:
----- 0/3 [h11]
      Successfully uninstalled httpcore-0.13.3
----- 0/3 [h11]
----- 1/3 [httpcore]
----- 1/3 [httpcore]
----- 1/3 [httpcore]
Attempting uninstall: httpx
----- 1/3 [httpcore]
    Found existing installation: httpx 0.19.0
----- 1/3 [httpcore]
    Uninstalling httpx-0.19.0:

```

```

----- 1/3 [httpcore]
Successfully uninstalled httpx-0.19.0
----- 1/3 [httpcore]
----- 2/3 [httpx]
----- 2/3 [httpx]
----- 3/3 [httpx]

```

Successfully installed h11-0.16.0 httpcore-1.0.9 httpx-0.28.1
Note: you may need to restart the kernel to use updated packages.

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.
googletrans 3.1.0a0 requires httpx==0.13.3, but you have httpx 0.28.1 which is incompatible.

```

import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

```

```

url=r"C:\Users\Dhanwantari Devre\VS CODE PROJECTS\0.TESTING\2...EDA
LLM\titanic_dataset_final.csv"
df=pd.read_csv(url)
df

```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
..	
886	887	0	2	
887	888	1	1	
888	889	0	3	
889	890	1	1	
890	891	0	3	

		Name	Sex	Age
SibSp	\			
0		Braund, Mr. Owen Harris	male	22.0
1				
1		Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0
1				
2		Heikkinen, Miss. Laina	female	26.0
0				
3		Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0
1				
4		Allen, Mr. William Henry	male	35.0
0				

```

..          ...
...
886          Montvila, Rev. Juozas      male  27.0
0
887          Graham, Miss. Margaret Edith  female  19.0
0
888          Johnston, Miss. Catherine Helen "Carrie"  female  NaN
1
889          Behr, Mr. Karl Howell      male  26.0
0
890          Dooley, Mr. Patrick        male  32.0
0

```

```

      Parch      Ticket      Fare Cabin Embarked
0         0      A/5 21171    7.2500   NaN      S
1         0      PC 17599   71.2833   C85      C
2         0  STON/O2. 3101282    7.9250   NaN      S
3         0      113803   53.1000  C123      S
4         0      373450    8.0500   NaN      S
..      ...
886        0      211536   13.0000   NaN      S
887        0      112053   30.0000   B42      S
888        2      W./C. 6607   23.4500   NaN      S
889        0      111369   30.0000  C148      C
890        0      370376    7.7500   NaN      Q

```

[891 rows x 12 columns]

```
print(df.describe())
```

```

count  PassengerId  Survived  Pclass     Age     SibSp  \
mean    446.000000    0.383838    2.308642  29.699118    0.523008
std    257.353842    0.486592    0.836071  14.526497    1.102743
min      1.000000    0.000000    1.000000    0.420000    0.000000
25%    223.500000    0.000000    2.000000   20.125000    0.000000
50%    446.000000    0.000000    3.000000   28.000000    0.000000
75%    668.500000    1.000000    3.000000   38.000000    1.000000
max    891.000000    1.000000    3.000000   80.000000    8.000000

```

```

count     Parch      Fare
mean     0.381594   32.204208
std     0.806057   49.693429
min      0.000000    0.000000
25%      0.000000    7.910400
50%      0.000000   14.454200
75%      0.000000   31.000000
max      6.000000  512.329200

```

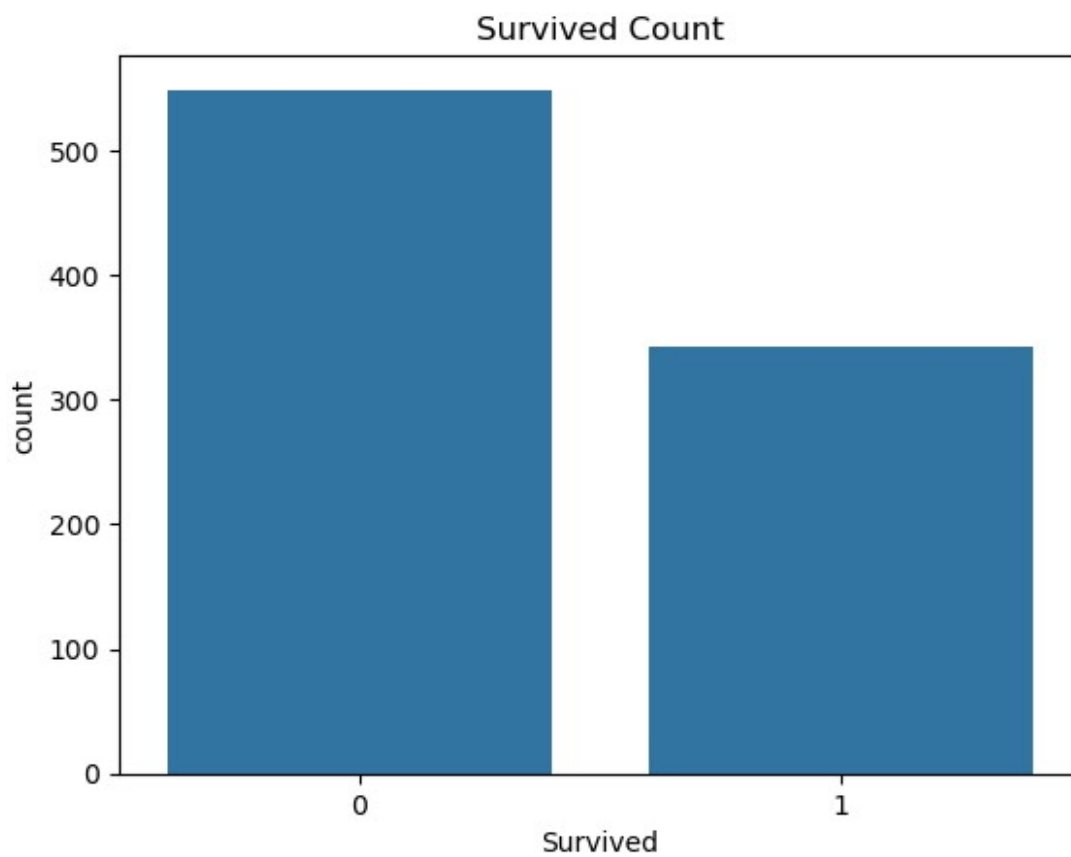
```
print("Missing Valuees = \n",df.isnull().sum())
```

```
Missing Valuees =  
PassengerId      0  
Survived          0  
Pclass           0  
Name             0  
Sex              0  
Age             177  
SibSp            0  
Parch            0  
Ticket           0  
Fare             0  
Cabin           687  
Embarked         2  
dtype: int64
```

```
sns.countplot(x='Survived',data=df)
```

```
plt.title("Survived Count")
```

```
Text(0.5, 1.0, 'Survived Count')
```



```

import ollama
def generate_insights(df_summary):
    prompt = f"Analyze the dataset summary and provide insights:\n\n{df_summary}"
    response = ollama.chat(model="mistral", messages=[{"role": "user", "content": prompt}])
    return response['message']['content']
# Generate AI Insights
summary = df.describe().to_string()
insights = generate_insights(summary)
print("\n AI-Generated Insights:\n", insights)

```

AI-Generated Insights:

The dataset is the Titanic Passenger data, which contains information about passengers who boarded the RMS Titanic on its maiden voyage from Southampton to New York in 1912. Here are some insights based on the provided summary:

1. There were a total of 891 passengers, with every passenger having an entry for each variable.
2. The average (mean) value for `Survived` is 0.383838, indicating that slightly less than half of the passengers survived.
3. Passengers belonged to different classes: 1st class (Pclass=1), 2nd class (Pclass=2), and 3rd class (Pclass=3). The mean Pclass is 2.308642, suggesting that most of the passengers were from 2nd class.
4. The average age is 29.699118 years, with a standard deviation of 14.526497, indicating a relatively wide range in ages.
5. On average, each passenger had 0.523 siblings (SibSp) and 0.381594 parents/guardians (Parch) traveling with them.
6. The average fare paid by passengers was 32.204208, with a standard deviation of 49.693429, suggesting significant variation in ticket prices.
7. The youngest passenger was 0.42 years old, while the oldest was 80 years old. Approximately 25% and 50% of passengers were below or equal to 20.125 years old and 28 years old, respectively. Around 75% of passengers were below or equal to 38 years old.

These insights offer a general overview of the data, but further analysis can be done to understand factors that influenced survival rates, ticket prices, and other aspects related to the Titanic disaster.

```
import gradio as gr
```

```
def greet(name):  
    return "Hello " + name + "!"  
  
demo = gr.Interface(fn=greet, inputs="text", outputs="text")  
demo.launch()  
  
* Running on local URL:  http://127.0.0.1:7860  
  
To create a public link, set `share=True` in `launch()`.  
<IPython.core.display.HTML object>  
  
import gradio as gr  
  
def eda_analysis(file):  
    df = pd.read_csv(file.name)  
    summary = df.describe().to_string()  
    insights = generate_insights(summary)  
    return insights  
  
# Create Web Interface  
demo = gr.Interface(fn=eda_analysis, inputs="file", outputs="text",  
    title="AI-Powered EDA with Mistral")  
  
# Launch App  
demo.launch(share=True)  # Use share=True for Google Colab  
  
* Running on local URL:  http://127.0.0.1:7861  
  
Could not create share link. Please check your internet connection or  
our status page: https://status.gradio.app.  
<IPython.core.display.HTML object>  
  
Created dataset file at: .gradio\flagged\dataset2.csv
```