

Fortune Global CEOs - Data Analytics

Import Libraries & Data

```
In [ ]: # Import the libraries we'll need
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [ ]: # Upload from your computer
from google.colab import files
print("Please upload the Fortune Global CEOs.xlsx file:")
uploaded = files.upload()
```

Please upload the Fortune Global CEOs.xlsx file:

No file chosen

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving Fortune Global CEOs.xlsx to Fortune Global CEOs.xlsx

```
In [ ]: df = pd.read_excel('Fortune Global CEOs.xlsx')
```

Initial Data Exploration

```
In [ ]: # Display basic information about the dataset
```

```
In [ ]: # Get summary statistics
```

```
In [ ]: # Check for missing values in each column
```

Data Cleaning and Preparation

```
In [ ]: df_clean = df.copy()
```

```
In [ ]: # Create a new column for CEO tenure
# Note: Some CEOs are still active, so we calculate based on current age
```

```
In [ ]: # For CEOs still active (where 'Age when CEO Tenure Finished' is NaN)
```

```
In [ ]: # Create a column for profit margin (%)
```

```
In [ ]: # Display dataframe
```

Exploratory Data Analysis (EDA)

```
In [ ]: # 1. Distribution of CEO ages
```

```
In [ ]: # 2. Industry distribution
```

```
In [ ]: # 3. Gender distribution
```

```
In [ ]: # 4. Salary distribution
```

```
In [ ]: # 5. Group by operations  
# Average salary by industry
```

```
In [ ]: # 6. Education impact on company performance
```

Business Implication Questions

```
In [ ]: # Question 1: Is there a relationship between CEO age at appointment and company pr
```

```
In [ ]: # Question 2: How does gender distribution vary across industries?
```

```
In [ ]: # Question 3: Top 5 highest-paid CEOs and their companies' profit margins
```

```
In [ ]:
```