

Salary Prediction System

Data Science - Final Project Presentation



Group Members

Eisha Tir Raazia

17k-3730

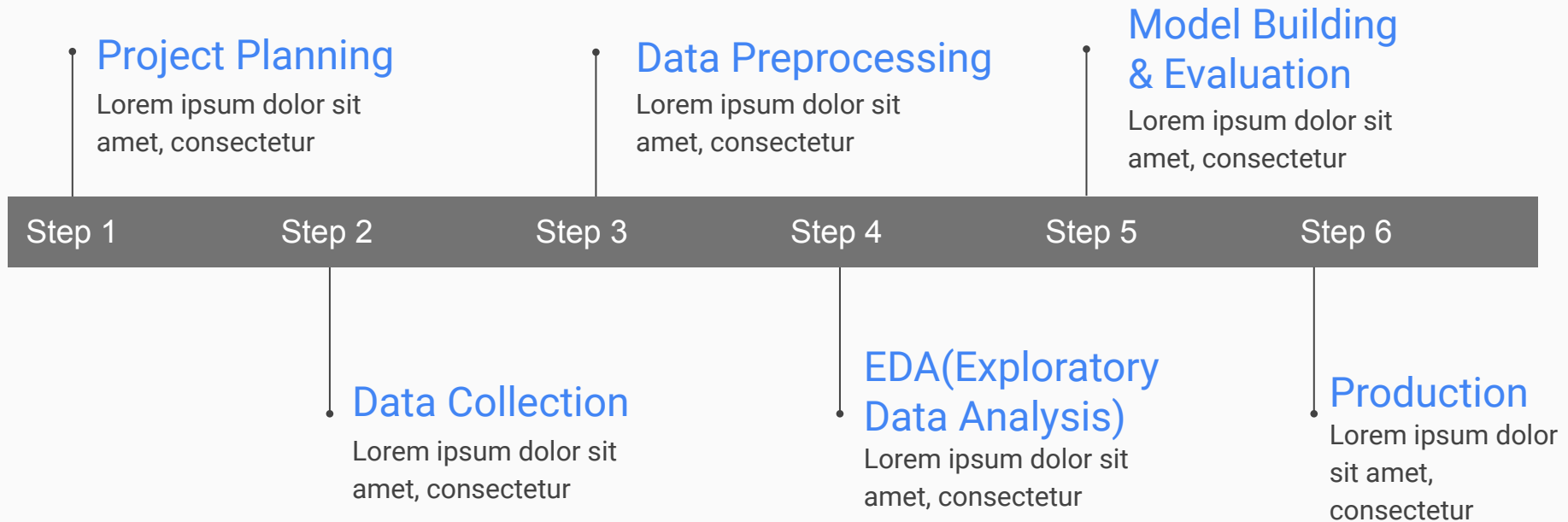
Maham Tariq

17k-3658

Methodology and Steps Involved



Steps Involved





Project Planning

Explored the available data sources and different websites.

Looked at existing work and tried to fit Data Science and ML use cases on available data.



Data Collection

Data Source:

Glassdoor.com

Wrote a Selenium based web scraper to extract Companies' data from the site by inputting various job roles.

Challenges we ran into while scraping

Sign Up dialog pop-up on webpage, at random times that was interrupting Execution.

Patience for exceptions when browser stops fetching data against a role. (Added max patience of 50)

Crashes before bulk updating the data into csv.



Data Cleaning and Preprocessing

Data Source:

Glassdoor.com

Wrote a Selenium based web scraper to extract Companies' data from the site by inputting various job roles.

Data Cleaning and Steps

Sign Up dialog pop-up on webpage, at random times that was interrupting Execution.

Patience for exceptions when browser stops fetching data against a role. (Added max patience of 50)

Crashes before bulk updating the data into csv.



Exploratory Data Analysis(EDA) & Visualizations

Data Source:

Glassdoor.com

Wrote a Selenium based web scraper to extract Companies' data from the site by inputting various job roles.

Data Cleaning and Steps

Sign Up dialog pop-up on webpage, at random times that was interrupting Execution.

Patience for exceptions when browser stops fetching data against a role. (Added max patience of 50)

Crashes before bulk updating the data into csv.



Model Building and Evaluation

Data Source:

Glassdoor.com

Wrote a Selenium based web scraper to extract Companies' data from the site by inputting various job roles.

Data Cleaning and Steps

Sign Up dialog pop-up on webpage, at random times that was interrupting Execution.

Patience for exceptions when browser stops fetching data against a role. (Added max patience of 50)

Crashes before bulk updating the data into csv.



Bringing Model to Production

Data Source:

Glassdoor.com

Wrote a Selenium based web scraper to extract Companies' data from the site by inputting various job roles.

How it works

Step 1

Lorem ipsum dolor sit
amet, consectetur



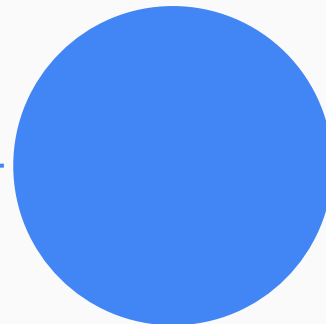
Step 2

Lorem ipsum dolor sit
amet, consectetur



Step 3

Lorem ipsum dolor sit
amet, consectetur

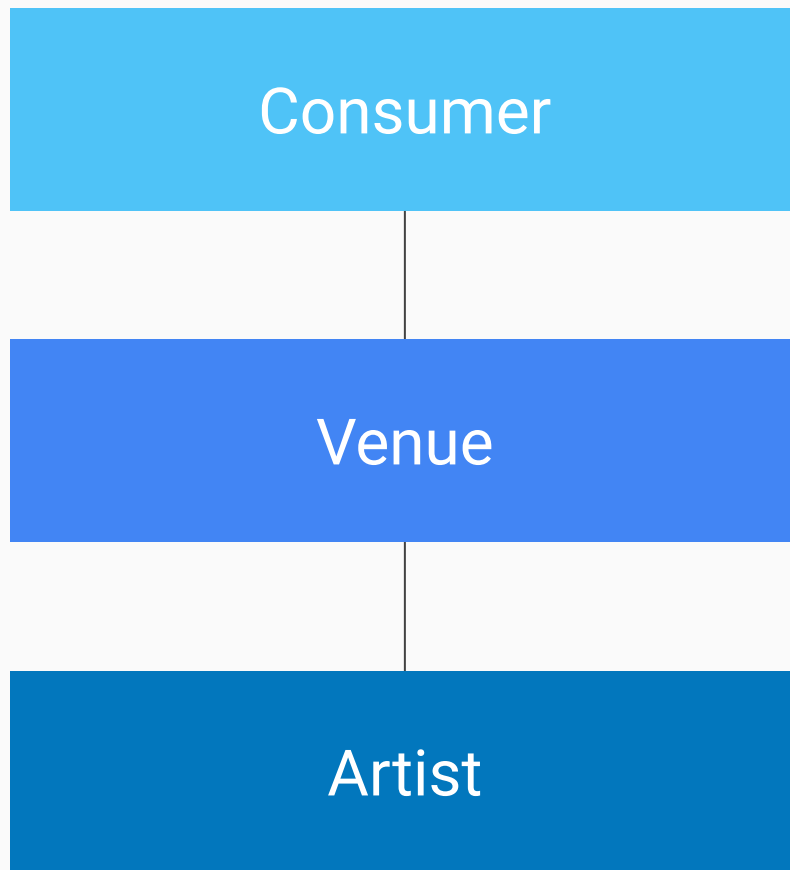


An aerial photograph of the New York City skyline at dusk. The sky is a mix of dark blue and orange, with scattered clouds. The city is densely packed with skyscrapers, many of which are illuminated with their interior lights. The Empire State Building is prominent in the center, with its top lit in red and green. The Hudson River is visible on the right side of the image. The text "Final Predictions" is overlaid in a large, white, sans-serif font in the center of the image.

Final Predictions

Revenue model

Lorem ipsum dolor sit amet,
consectetur adipiscing elit,
sed do eiusmod tempor
incididunt



Conclusion:

