

ITCS-6100 Big Data for Computational Advantage

Group -18

Project Deliverable- 1

1. Team

a) Team Members

Manasa Avula - 801307493
Nikhita Sai Boyidapu - 801327682
Srikar Chamarthy - 801317299
Rachana Gullipalli - 801311637
Aravind Pabbisetty - 801274519

b) Communication Plan to include project artifact repository.

- The group meetings will be held weekly twice virtually on Google Meet. Every group member agrees to attend the meetings on time. These meetings are conducted based on the project deliverables. Further group members can also discuss and chose a time if an in-person meeting is required in some cases.
- Members agreed to respect one another and share and discuss their ideas.
- Members can share their findings with others through Emails, Group Shared Folder

Folder Link:

https://drive.google.com/drive/folders/1ObKCKcFXyW_kZzUwfqe0GMhXDjKxhRF-T?usp=share_link

GitHub repository URL:

<https://github.com/aravindpabbisetty/BigDataGroup18>

Choice 1

2. Selection of data to analyse from Kaggle:

Dataset Link:

<https://www.kaggle.com/datasets/ricardoaugas/salary-transparency-dataset-2022>

3. Business Problem or Opportunity, Domain Knowledge

This Salary Transparency dataset consists of salary information of employees working in various companies along with other information such as Education, Industry, Job Title, Experience, etc.

The business opportunity to identify from this dataset is that it helps the organizations to understand the salaries provided by other companies and can help them to understand whether they are underpaying or overpaying the employees. This can also help them in increasing the employee retention.

4. Research Objectives and Questions

- 1) Which industry is having high demand and salary?
- 2) What is the average salary for each job title?
- 3) Which profession will have high demand in next 5 years?
- 4) What is the average education required to get into Finance sector?
- 5) What is the most preferred profession between 18-21 age group?

Choice 2

2. Selection of data to analyse from Kaggle:

Dataset Link:

<https://www.kaggle.com/datasets/kushleshkumar/cornell-car-rental-dataset?select=CarRentalDataV1.csv>

3. Business Problem or Opportunity, Domain Knowledge

This Car Rental dataset consists rental cars information such as make, model type, fuel types, location, etc. The business opportunity that we can identify is this dataset helps to analyse customer preferences in rental cars. This helps the businesses to improve their car services by upgrading their models, meet their customer demands, and provide better services to satisfy their customers.

4. Research Objectives and Questions

- 1) What are the most popular car models and brands among customers?
- 2) Which car model and make will be on demand in next 5 years?
- 3) Which location is having high demand for rental cars?
- 4) Which fuel type of cars does the customers preferred to rent in Atlanta?
- 5) Which fuel type cars will be on demand in next 3 years?