**Jolina Minezes**

Location| Contact| Email| [Linkedin](https://www.linkedin.com/in/jolina-minezes-0122b8247/%20) | [Github](https://github.com/Jolina1205)

**EDUCATION**

|  |
| --- |
| **NSOMASA, NMIMS, Mumbai** *Jul 2022 – May 2024*  *MSc Statistics & Data Science*  CGPA – |
| **D.G.Ruparel College of Arts , Science & Commerce** *Jul 2019 – May 2022*  *BSc Statistics*  CGPA – |

**WORK EXPERIENCE**

|  |  |
| --- | --- |
| **BINARY DATA LABS – Analyst Intern** | *Feb 2024 – Jun 2024* |

* Extracted and processed 100+ travel industry financial data from Yahoo Finance.
* Productionized data on Google Cloud Platform (GCP) for analysis in BigQuery.
* Implemented data cleaning by integrating Great Expectations for automated data quality checks and validation tests.
* Collaborated on financial analysis for 100+ travel industry companies.
* Currently contributing to the development of a Large Language Model (LLM) for financial analysis of travel companies' annual reports using the GPT-3.5 OpenAI model.
* Leveraged Scrapy for efficient web crawling and data extraction.
* Employed Retrieval-Augmented Generation (RAG) technique and Langchain, LLM on Hugging face.
* Manually prepared summaries of annual reports to validate the accuracy of the LLM model.
* Contributing to the development of a financial analysis tool encompassing dashboarding and automation for multiple industries.

**PROJECTS**

|  |  |
| --- | --- |
| **Automatic Number Plate Recognition Systems** | *Feb 2023 – April 2023* |

* Trained a Convolutional Neural Network (CNN) on a dataset of 484 Indian car images, achieving accurate license plate recognition.
* Performed image processing & object detection using OpenCV & character segmentation & recognition using TensorFlow.
* Applied EasyOCR for digital character recognition, resulting in **87% accuracy**.

|  |  |
| --- | --- |
| **Diamond Price Prediction** | *March 2023 – Apr 2023* |

* Conducted data preprocessing and exploratory data analysis (EDA) on a historical Diamond dataset comprising 53,940 entries with 10 features, leading to the development of a predictive model for diamond price prediction.
* Used python libraries: Pandas, numpy, seaborn, sklearn.
* Compared 3 Statistical models; linear regression, Decision tree regressor and Random Forest regressor.
* Achieved a minimal **1.25% mean squared error** with the Random Forest Regressor, providing the best model fit attaining

**76% accuracy**.

|  |  |
| --- | --- |
| **Restaurant Sales Data Analysis** | *Sept 2022 – Nov 2022* |

* Conducted comprehensive restaurant sales data analysis, employing Excel & SQL techniques for data preprocessing and visualization.
* Developed an Excel dashboard showcasing key performance indicators, enabling stakeholders to identify actionable insights for driving business growth and achieving around a 15- 20% increase in customer retention.
* Employed advanced SQL queries, including subqueries, RANK, and DENSE\_RANK, to identify top-selling cuisine, peak sales periods, and assess customer preferences.

**SKILLS & INTERESTS**

**Hard Skills:** Google Cloud Platform**|** Python | SQL | Tableau | Github | MS Excel | R | Data Extraction | EDA & Visualization | Report Summarization | Statistical & Financial Analysis | Predictive Modelling | Machine Learning | NLP

**Interests:** Finance | Traveling | Pencil Sketching

**EXTRA CURRICULAR ACTIVITIES**

**Awards & Achievements**

* Developed SQL proficiency by earning a prestigious Badge on Hacker-Rank, successfully completing 50+ coding challenges.

**CERTIFICATIONS**

* Business Analytics with Excel – Simplilearn
* Introduction to SQL - Simplilearn
* Virtual Experience of Accenture with Excel (Data Analytics and Visualization) by Forage
* Virtual Experience of Tata with Tableau Prep Builder (Data Visualization) by Forage