

Lesley-w19 / Phase3-Project Public

Billionaire Dataset exploration in this Project 3

0 stars 0 forks Activity


Star

Notifications

- <> Code
- Issues
- Pull requests
- Actions
- Projects
- Security
- Insights

main

Go to file


 Lesley-w19 Update on the presentation pdf ... 1 minute ago 12

[View code](#)

☰ README.md

PHASE 3 PROJECT [↗](#)

Stepping into Billionaire Territory!! [↗](#)



Project Overview [↗](#)

This project requires the use Classification Model to generate insights for a given Agency. This agency provides all sorts of social support in the different industries, mostly volunteer work. In a bid to increase their profits for the next financial year, their have been looking through different datasets that will enable them make a choice of industriy investment.

Business Understanding [↗](#)



Stakeholder Audience [↗](#)

Our Stakeholder are Soar High Research Agency, a non-governmental organization that deals with Social Development. They are looking for an industry to invest in that will yield profit that will be able to fund their various Activities (eg. school sponsorships).

Based on their source question, we are required to provide insights on the wealth distribution status and profitable industries they are to venture in. Also, as the kind of industries they are able to encourage their students to take.

Business Objectives [↗](#)

The aim of this project is to analyze the provided dataset on King Country and eventually answer to the Stakeholders questions. With this in mind, the objectives of this project include: The objectives of this project based on the dataset chosen is to find out:

1. The Wealth Status of Billionaires (whether self-made or not)
2. What industry/sources is more inclined to produce billionaires?
3. Demographic analysis of billionaires (age, gender, country)
4. Provide classification to the wealth status of billionaires based on the features.
5. Provide insights into which industry are likely to produce billionaires in future (logistic regression)g

Data Understanding [↗](#)

For this project, we will use the Billionaires Statistics dataset under the file path `data\...` This folder contains historical data of billionaire information captured in a northwestern country, which we will modify for the purpose of our analysis.



The data folder includes the following file:

1. Billionaires Statistics Dataset.csv - provides the billionaire dataset which will be used for our analysis

The dataset represented has the following data columns as some of the header categories:

rank : The ranking of the billionaire in terms of wealth. finalWorth : The final net worth of the billionaire in U.S. dollars. category : The category or industry in which the billionaire's business operates. personName : The full name of the billionaire. age : The age of the billionaire. country : The country in which the billionaire resides. city : The city in which the billionaire resides. source : The source of the billionaire's wealth. industries : The industries associated with the billionaire's business interests. countryOfCitizenship : The country of citizenship of the billionaire.

Visualization [↗](#)

Modelling [↗](#)

Evaluation [↗](#)

Recommendation [↗](#)

Conclusions [↗](#)

Releases

No releases published

Packages

No packages published

Languages

● Jupyter Notebook 100.0%