

Life Expectancy Analysis

CRISP-DM

Understanding Business

Assessing Current Situation

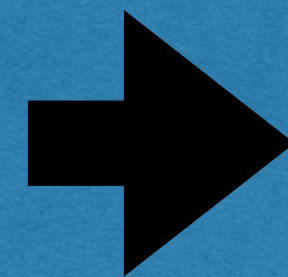
Inventory of Resources

- Data Analyst , Python Programmer (Personnel)
- LE-File Description.docs, education.csv, crime.csv, income.xlsx, area.csv, lifeExpectancy.csv, region.txt (Data)
- MacOS Big Sur , 8GB RAM, Mac M1 chip (Computing Resources)
- Anaconda Distribution, Python 3, Numpy, Pandas, Matplotlib, Seaborn, Statsmodels, Terminal, Git, Jupyter Notebooks, Keynote (Software)

Assessing Current Situation

Schedule of Completion of Project (32 Days)

Tasks Schedule



Project Steps	Duration (in Days)	Start Date	End Date
CRISP-DM (Ongoing)	32	1.7.2022	31.7.2022
Data Collection & Cleaning	9	2.7.2022	5.7.2022
Data Exploration	8	6.7.2022	16.11.2022
Unusual Observations	12	17.7.2022	20.7.2022
Regression Analysis	7	21.7.2022	27.7.2022
Results & Summary	6	28.7.2022	31.7.2022

Assessing Current Situation

Project Risks & Potential Solutions

- Over generalisation to other population outside America
- Lack of clarity
- Solution would be to restrict our understanding to limited area and Planning effectively

Business & Data Mining Goals

Primary Business Objective

Our Primary Objective would be to find factors related to life expectancy America (Business) by running Regression Analysis (Data Mining) on the data to find factors related to life expectancy by finding out positively correlated factors (Measure of Success).

Business & Data Mining Goals

Secondary Business Objective

Our Secondary Objective would be to know more about distribution of certain demographic variables in US population (Business) by analysing data to answer a series of questions (Data Mining) and finding out and plotting results (Measure of Success).

Project Plan & Tools Required

Project Steps	Duration (in Days)	Start Date	End Date	Tools & Techniques
CRISP-DM (Ongoing)	32	1.7.2022	31.7.2022	Understanding Analysis Life
Data Collection & Cleaning	4	2.7.2022	5.7.2022	Numpy, Pandas, Data Preparation
Data Exploration	11	6.7.2022	16.11.2022	Data Visualization, Seaborn, Plotting
Unusual Observations	4	17.7.2022	20.7.2022	Outlier Detection
Regression Analysis	7	21.7.2022	27.7.2022	Regression, Statsmodels
Results & Summary	4	28.7.2022	31.7.2022	Plots, Summary, Report