

# Life Expectancy Analysis

## CRISP-DM

# Understanding Business

# Assessing Current Situation

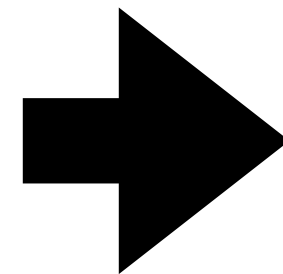
## Inventory of Resources

- 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 (26 Days)

Tasks Schedule



Project Steps	Duration (in Days)	Start Date	End Date
<b>CRISP-DM (Ongoing)</b>	32	21.10.2021	3.12.2021
<b>Data Collection &amp;</b>	9	22.10.2021	30.10.2021
<b>Data Exploration</b>	8	1.11.2021	8.11.2021
<b>Unusual Observations</b>	12	9.11.2021	20.11.2021
<b>Regression Analysis</b>	7	21.11.2021	27.11.2021
<b>Results &amp; Summary</b>	6	28.11.2021	3.12.2021

# 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
<b>CRISP-DM (Ongoing)</b>	32	21.10.2021	3.12.2021	Understanding Analysis Life
<b>Data Collection &amp; Cleaning</b>	9	22.10.2021	30.10.2021	Numpy, Pandas, Data Preparation
<b>Data Exploration</b>	8	1.11.2021	8.11.2021	Data Visualization, Seaborn, Plotting
<b>Unusual Observations</b>	12	9.11.2021	20.11.2021	Outlier Detection
<b>Regression Analysis</b>	7	21.11.2021	27.11.2021	Regression, Statsmodels
<b>Results &amp; Summary</b>	6	28.11.2021	3.12.2021	Plots, Summary, Report