Predicting Life Expectancy from Country Demographics

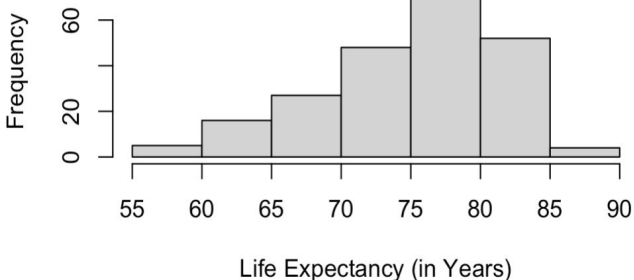**Stat 418 - Tools in Data Science**
Gaohong Chen

Data Collection

- Data Source
  - U.S. Census Bureau International Data Base (IDB)
  - Accessed public API
- Access & Processing
  - Extract raw data for the year 2024, including over 220 countries and territories
  - Convert the raw JSON response and transformed into tabular dataset
- Demographic Variables
  - POP: Population
  - GR: Growth Rate
  - POP_DENS: Population Density
  - TFR: Total Fertility Rate
  - IMR: Infant Mortality Rate
  - CDR: Crude Death Rate
- Target Variable
  - E_0: Life Expectancy

Exploratory Data Analysis

Distribution of Life Expectancy at Birth

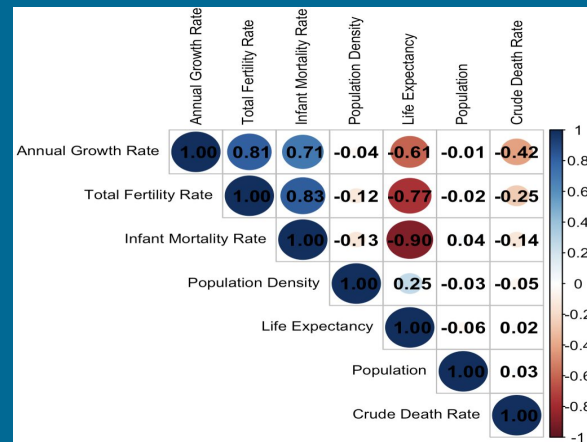Infant Mortality vs. Life Expectancy

- Most countries have life expectancy between **70 and 85 years**
- A few countries have life expectancy **below 65 years**
- The chart is slightly **skewed to the left**

- Strong negative relationship
- Countries with **low infant mortality** tend to have **longer life expectancy**

- Fertility rate has a strong negative correlation with life expectancy
- Growth rate shows a moderate negative correlation
- Growth rate and fertility rate are strongly positively correlated

Modeling and Flask API

- Modeling
  - Multiple Linear regression using Scikit-Learn in Python
- API Deployment
  - Flask API created and containerized with Docker
  - Hosted on Google Cloud Run
  - Accepts JSON input, returns predicted life expectancy
- Request Example
  - Curl command to test live API
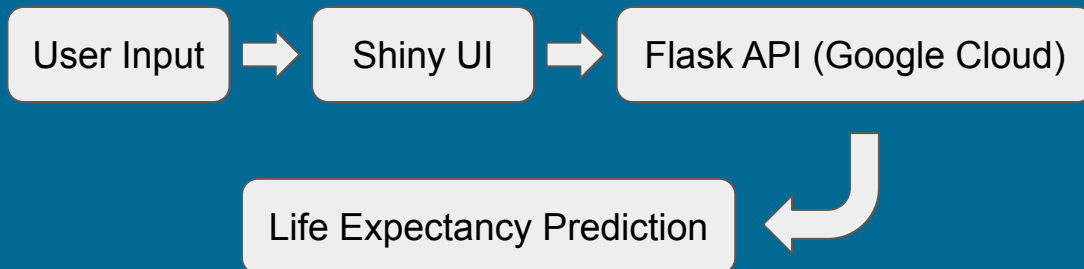 - Response shows numeric prediction

```
curl -X POST https://final-project-flask-app-693104809772.us-central1.run.app/predict_price \
  -H "Content-Type: application/json" \
  -d '{
    "Population": 1000000,
    "Annual Growth Rate": 1.2,
    "Population Density": 500,
    "Total Fertility Rate": 2.3,
    "Infant Mortality Rate": 20,
    "Crude Death Rate": 8
}'

{"predicted life expectancy":74.78086113064549}
```

Shiny App

- The Shiny app interacts with a Flask API hosted on Google Cloud Run
- Allows users to input demographic data for prediction
- Link to the app
  - https://gaohongc.shinyapps.io/life_expectancy/

User Input ➡ Shiny UI ➡ Flask API (Google Cloud)

Life Expectancy Prediction ⬅

📈 Predict Life Expectancy by Country

Predicted Life Expectancy:

Enter demographic data below to estimate life expectancy. **Note:** Country selection is for reference only. Predictions are based on the input demographic values.

🌐 *Powered by Flask API hosted on Google Cloud Run*

Select Country:

Afghanistan ▾

Population:

1000000

Annual Growth Rate (%):

1.2

Population Density:

500

Total Fertility Rate:

2.3

Infant Mortality Rate:

20

Crude Death Rate:

8

Predict