**Secondary Datasets Included for Our Analysis and Vizualisation**

**Glucose Levels**

<https://www.cdc.gov/diabetes/basics/getting-tested.html>

resources folder file name: glucose\_test.csv

ETL folder file name: glucose\_pks.csv

The dataframe contains two columns. The first contains three different ranges of fasting blood sugar levels and the corresponding results or descriptions for each of the ranges.

The corresponding results or descriptions will be integrated, after the glucose column, in the primary dataset used for this analysis and visualization.

### Fasting Blood Sugar Test(as described by the CDC)

This measures your blood sugar after an overnight fast (not eating). A fasting blood sugar level of 99 mg/dL or lower is normal, 100 to 125 mg/dL indicates you have prediabetes, and 126 mg/dL or higher indicates you have diabetes.

**Blood Pressure Dataset**

https://www.heart.org/en/health-topics/high-blood-pressure/understanding-blood-pressure-readings

resources folder file name: blood\_pressure.csv

ETL folder file name: bp\_pks.csv

The dataframe contains four columns. The first contains the five categories a person might fall into depending on their blood pressure reading. The second column contains the range of values for systolic BP and the last column contains the values for diastolic BP.

The corresponding blood pressure categories will be integrated, after the systolic and diastolic BP columns, in the primary dataset used for this analysis and visualization.

The paragraphs below describe what the blood pressure measurement means; It is taken from the website of the American Heart Association.

Your blood pressure is recorded as two numbers:

**Systolic blood pressure** (the first number) – indicates how much pressure your blood is exerting against your artery walls when the heart beats.

**Diastolic blood pressure** (the second number) – indicates how much pressure your blood is exerting against your artery walls while the heart is resting between beats.

Which number is more important?

Typically, more attention is given to systolic blood pressure (the first number) as a major risk factor for cardiovascular disease for people over 50. In most people, systolic blood pressure rises steadily with age due to the increasing stiffness of large arteries, long-term buildup of plaque and an increased incidence of cardiac and vascular disease.

However, either an elevated systolic or an elevated diastolic blood pressure reading may be used to make a diagnosis of high blood pressure. According to recent studies, the risk of death from ischemic heart disease and stroke doubles with every 20 mm Hg systolic or 10 mm Hg diastolic increase among people from age 40 to 89.

**BMI Datset**

https://www.cdc.gov/healthyweight/assessing/bmi/adult\_bmi/index.html

resources folder file name: bmi.csv

ETL folder file name: bp\_pks.csv

The dataframe contains two columns. The first column contains a range of values for BMI and the second column contains the corresponding descriptions used by health professionals.

The corresponding bmi descriptions will be integrated, after the bmi column, in the primary dataset used for this analysis and visualization.

The paragraphs below are from the CDC’s website and it describes how a person’s BMI is computed, and briefly eplains how it may affect one’s health.

Body mass index (BMI) is a person’s weight in kilograms divided by the square of height in meters. BMI is an inexpensive and easy screening method for weight category—underweight, healthy weight, overweight, and obesity.

BMI does not measure body fat directly, but BMI is moderately correlated with more direct measures of body fat 1,2,3. Furthermore, BMI appears to be as strongly correlated with various metabolic and disease outcome as are these more direct measures of body fatness 4,5,6,7,8,9.

**Total Cholesterol**

Source: https://www.mayoclinic.org/tests-procedures/cholesterol-test/about/pac-20384601

resources folder file name: total\_cholesterol.csv

ETL folder file name: totchol\_pks.csv

The dataframe contains three columns. The first column contains a range of values of total cholesterol measured in mg/dL, the second column also contains total cholesterol but instead is measured in mmol/dL and the last column contains the corresponding descriptions used by health professionals to describe total cholesterol.

This paragraph explains what cholesterol and total cholesterol is.

Total cholesterol is the total amount of cholesterol in your blood. Your total cholesterol includes low-density lipoprotein (LDL, or “bad”) cholesterol and high-density lipoprotein (HDL, or “good”) cholesterol. Cholesterol is a waxy, fat-like substance found in every cell in your body.(source: https://www.verywellhealth.com/what-is-a-total-cholesterol-level-698073)