Dr. lamba’s Minute ClinicS

Team Members:

|  |
| --- |
| Vin Dixit |
| Dinesh Lamba |
| Kendall Gouldthorpe |
| Wesley Stone |

Customer:

Dr. Lamba’s Minute Clinic is a Independent Practitioners Association (IPA) of over 1,100 clinics across the United States, offering convenient and high-quality walk-in medical care options. They hire Nurse Practitioners and Physicians’ Assistants who are able to provide medical care to patients with minor illnesses and injuries including health screenings and monitoring services, vaccinations and injections as well as wellness and physicals.

Customer Proposal:

Dr. Lamba’s Minute Clinic is looking to provide additional information to diabetic patients that come in for any screening and monitoring that relate to their diabetic conditions. Dr. Lamba’s clinic wants to find if there is any correlation between body vitals such as Body Mass Indices (BMI), Age and medical conditons like diabetes and heart attack. They would also like to investigate if there exists any interdependency between diabetes and heart attack.

**Hypothesis:**

1. There is a correlation between body vitals such as age, BMI and diabetes
2. There is a correlation between body vitals such as age, BMI and heart attack
3. There is correlation between patients with diabetes and propensity for heart attack

**Null:**

1. There is no correlation between body vitals such as age, BMI and diabetes
2. There is no correlation between body vitals such as age, BMI and heart attack
3. There is no correlation between patients with diabetes and propensity for heart attack

Data Sources:

**Data Set #1:**

Cardiovascular Disease Dataset

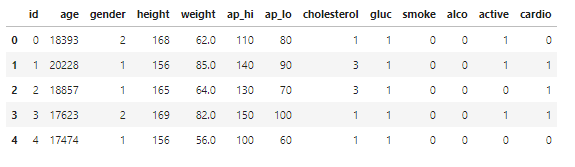
<https://www.kaggle.com/sulianova/cardiovascular-disease-dataset>

File Type: .csv

Rows: 70,000

Data includes:

1. Age | Objective Feature | age | int (days)
2. Height | Objective Feature | height | int (cm)
3. Weight | Objective Feature | weight | float (kg)
4. Gender | Objective Feature | gender | categorical code
5. Systolic blood pressure | Examination Feature | ap\_hi | int
6. Diastolic blood pressure | Examination Feature | ap\_lo | int
7. Cholesterol | Examination Feature | cholesterol | 1: normal, 2: above normal, 3: well above normal
8. Glucose | Examination Feature | gluc | 1: normal, 2: above normal, 3: well above normal
9. Smoking | Subjective Feature | smoke | binary
10. Alcohol intake | Subjective Feature | alco | binary
11. Physical activity | Subjective Feature | active | binary
12. Presence or absence of cardiovascular disease | binary



**Data Set #2:**

Diabetes

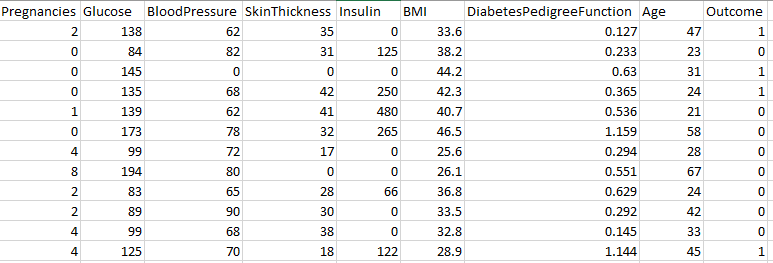
<https://www.kaggle.com/johndasilva/diabetes>

File Type: .csv

Rows: 2,000

Data includes:

1. Pregnancies | Pregnancies | int
2. Glucose | Glucose | int
3. Blood Pressure | BloodPressure | int
4. Skin Thickness | SkinThickness | int
5. Insulin | Insulin | int
6. BMI | BMI | float
7. Diabetes Pedigree Function | DiabetesPedigreeFunction | float
8. Age | Age | int
9. Outcome | Outcome | binary



Deliverables from GDLW Consultants:

1. Report describing methodology of data collection and data cleaning
2. Two cleaned datasets
3. Various views joining the datasets that would help client to analyze for given hypothesis