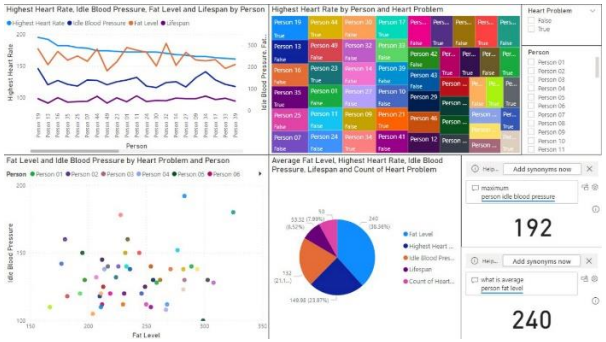

Link : https://app.powerbi.com/view?r=eyJrIjoiYzQ4NWQ1NjgtNTU4Mi00YjY0LWFjOGMtNWYzODA0ZDVkM2U4IiwidCI6IjBlMGRiMmFkLWM0MTYtNDdjNy04OGVjLWNlYWM0ZWU3Njc2NyIsImMiOjEwfQ%3D%3D&pageName=ReportSection

Part 1 : Group Details

Title : Heart Failure Prediction Dataset

Group Leader : Wan Muhammad Aiman bin Wan Muhamad Faisal (A21EC2037)

	Group Excel	Group Microsoft Power BI
Group Members	• Wan Muhammad Aiman bin Wan Muhamad Faisal (A21EC2037) • Muhammad Fahmi bin Roslee (A21EC0285)	• Aaron Tan (A21EC0152) • Afiq Fahmi bin Roslan (A21EC0153) • Muhammad Thoriq bin Kahairi (A21EC0096)
Tasks	• Give title of the report • Describe Trends of Data Analytics in Different Sectors and reflection of these trends • A narrative of the analytics. • Download one of the dataset from Kaggle (https://www.kaggle.com/datasets) • Combine and clean data, and prepare Excel spreadsheets • Import data into Microsoft Excel • Change variable names and labels • Check for missing data types and variables • Transform any variables that you would like to use in a different form • Discussion of the results for data visualization • Conclusion	• Describe Industrial Talk 7 related to Microsoft Power BI as discussed in the talk; what, when and where to use it; and the reflection from the talk • Perform data analytics using Microsoft Power BI (https://powerbi.microsoft.com/en-us/) • Interpret the data analytics • Discussion of the results for data visualization

Part 2 : Introduction

Trends of Data Analytics in Different Sectors

Data analysis is a set of strategies for extracting meaningful and valuable information from large and diverse sets of data acquired from many sources and ranging in size. There are four types of data analytics which are descriptive, diagnostic, predictive, and prescriptive. Descriptive analytic describes the events that occur across time, such as whether the current month's revenues are higher than the previous month's. Next, diagnostic analytic focuses on the cause of any event's occurrence It involves hypothesizing and involves a large and diversified dataset. Other than that, predictive analytics focuses on occurrences that are predicted to happen in the near future. Last but not least, prescriptive analytics implies a course of action.

By analysing heart failure prediction database, we can use this data as a supporting tool for doctors and physicians to predict if a patient is going to have a heart failure. Thus, they can advise their patient on what to do for them to avoid a heart failure and live a longer life based on the dataset value such as fat level, idle blood pressure and highest heart rate.

Industrial Talk 7

From the talk, Microsoft Power BI is a reporting tool that allows transformation of data into a more presentable and interactive one. The data is usually obtained from Microsoft Excel or the SQL server. The reason why people use Microsoft Power BI is the connection of data and allows other colleagues or friends that have access to interact with the data with ease through the cloud. Microsoft Power BI also introduces numerous ways to represent the data such as in a treemap, funnels and filled map which allows flexibility. This is mainly used in business companies to represent their sales with easy to see and access visuals for the relevant data.

These models are then sent to their organization to access and make optimal solutions based on the data. For example, from the talk, the workers or fellow colleagues are able to access very specific data like sales from a state of a country, type of categories and more just by clicking on them. This allows them to easily find the data that is required by them instead of having a rough time searching for the data especially if there are over hundreds of countries and categories within the model. Thus, this allows the company to make the right decisions with ease and focus on which part of the country or category to make changes for the better.

Narrative of the analytics

Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide.

Heart failure is a common event caused by CVDs and this dataset has 3 features that can be used to predict mortality by heart failure.

Part 3 : Data analytics interpretation for the chosen data

Variable transformation

Changed variable :

1. Age > Lifespan
2. RestingBP > Idle Blood Pressure
3. Cholesterol > Fat Level
4. MaxHR > Highest Heart Rate
5. HeartDisease > Heart Problem

Changed variable transform :

1. IBP (whole number > with 1 dp)
2. FL (whole number > with 2dp)

Discussion

Highest Heart Rate, Idle Blood Pressure, Fat Level and Lifespan by Person

Based on the highest heart rate, idle blood pressure, fat level and lifespan by person in the line chart graph, we can see that person 19 has the highest heart rate which is 195 and also the highest blood pressure which is 192 that leads to heart problem disease. Not only that, person 47 has the highest fat level which is 299 among the people who have heart problems. However, person 32 has the highest fat level which is 308 among all the people but did not get a heart problem. Furthermore, person 45 and person 46 have the highest lifespan, which is 69 but only person 46 has a heart problem. Tendency to get heart problems that are affected by lifespan and fat level are not 100% but aging can cause changes in the heart and blood vessels that may increase a person's risk of developing cardiovascular disease while with high cholesterol or fat level, you can develop fatty deposits in your blood vessels. Eventually, these deposits grow, making it difficult for enough blood to flow through your arteries. Sometimes, those deposits can break suddenly and form a clot that causes a heart attack or stroke. We can conclude that, high heart rate, high blood pressure, high fat level and high lifespan can lead to the increasing of the number of heart problems.

Highest Heart Rate by Person and Heart Problem

From the tree map, the highest heart rate for heart problem with value of True (meaning the person has a heart problem) is person 19 on the top left corner of the map with 195 Heart Rate. Meanwhile, the highest heart rate for heart problems with value of False (meaning the person does not have a heart problem) is person 13 which is below person 19 from the map with 192 Heart Rate. Furthermore, the lowest heart value is person 20 with 95 Heart Rate for True and person 3 with 98 Heart Rate for False. The tree map arranges the person with highest heart rate on the top left and the person with the lowest heart rate value on the bottom right. We can also see that as the graph descends to the bottom right, the surface area of the box of the person is decreasing showing that the heart value also decreases. From this representation of data, we can see that the highest heart rate of a person is not really a major cause of people having heart problems as even people with the lowest of the lowest heart rate still have heart problems.

Fat Level and Idle Blood Pressure by a Heart Problem and Person

Based on the fat level and idle blood pressure by a heart problem and person in the scatter plot graph, it can be seen person who has a heart problem with the highest blood pressure from person 19 which is 192 while the person 47 has the highest fat level which is 299 among people who have a heart problem. The data also shows a person with a heart problem has an idle blood pressure of above 140. Tendency to get heart problems is high if the person has high idle blood pressure above 140. The normal idle blood pressure for humans is from120mmHg to 130mmHg. Moreover, high-fat levels might cause the blood pressure of persons to be increased. Person with obesity causes the blood to coagulate and prevents oxygen to be transferred and can increase the number of heart diseases .We can conclude that high-fat level, and high idle blood pressure influence the number of heart problems. However, the result of the heart problem can be changed due to the age that needs to be considered as well.

Average Fat Level, highest Heart Rate, Idle Blood Pressure, Lifespan and Count of Heart Problem

According to the pie chart, the average fat level is 240, the average highest heart rate is

149.98, the average idle blood pressure is 32, the average lifespan is 53.32, and the average heart problem is 50. The average fat level consumes 38.96% of the pie chart, which is the largest data value, while the average heart problem consumes just 7.99% of the pie chart, which is the lowest data value.

Conclusion

Most cardiovascular diseases can be prevented by addressing behavioural risk factors such as tobacco use, unhealthy diet and obesity, physical inactivity, and harmful use of alcohol. People with cardiovascular disease need early detection and management wherein a machine learning model can be of significant help. By using machine learning model, we can have an early detection of people that have cardiovascular disease thus helping them manage their life to help them prevent getting a heart failure.

Reference

Emerging Trends: Data Analytic Slide.

Mr. Isma Redha (23 December 2021), “Talk 7: Introduction to Data Visualization (iCEP)”.

fedesoriano. (September 2021), ‘Heart Failure Prediction Dataset. Retrieved’ [Date Retrieved : 25 December 2021] from https://www.kaggle.com/fedesoriano/heart-failure-prediction.