

Washington State University
MIS 420 – Business Intelligence
Online

Mauricio Featherman, Ph.D.

T-SQL #4

Name: Nam Jun Lee

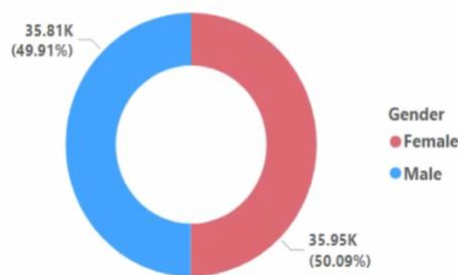
Student Number: 11606459

Introduction

This project focuses on discovering the relationship and insights between each variable using the health and potential coronary disease indicators in Project 3 conducted earlier. The goal is to use the array for deeper and more insightful analysis and to provide health clinics with professional reports that can guide patients to healthier and happier life.

Data Analysis

Distribution by Gender



Gender	ID
Female	35946
Male	35814
Total	71760

Through the left pie chart and above matrix, the total number of patients is 71,760 and the gender ratio of all patients is 49.91% for males and 50.09% for females, at a similar level.

Count of Treadmill Result by Age Group

Heart Health Specific ● Warning ● Normal ● Healthy ● Extremely Healthy ● At risk



Looking at the chart above, you can check the treadmill result values for each age group.

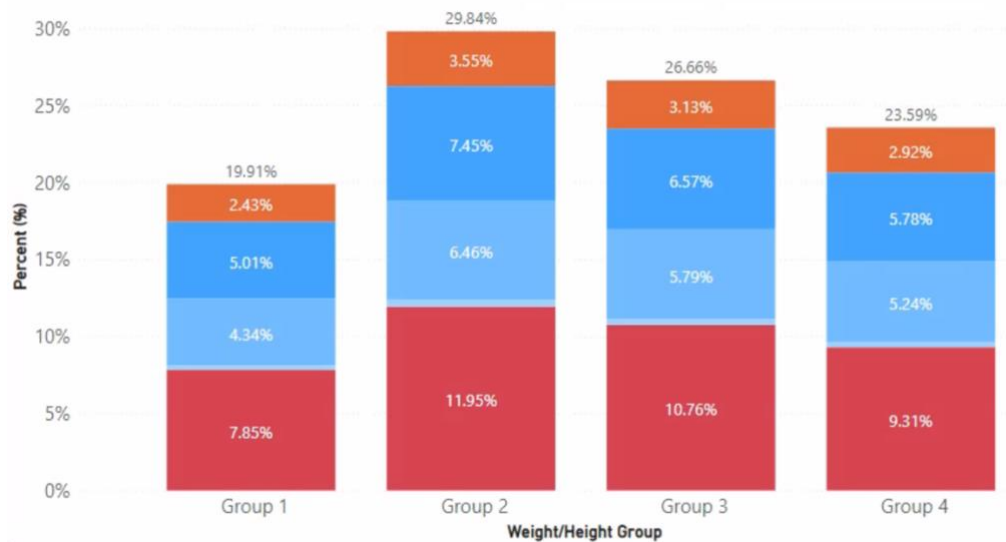
There are five labels here.

1. At risk: 100% reached or exceeded
2. Warning: 90% reached, 100% not reached
3. Normal: 70% reached 90% not reached
4. Healthy: 50% reached 70% not reached
5. Extremely Healthy: Reached only 50%

According to the results, there are many people in their 20s with better heart health than normal, and as the age group increases, fewer people with better heart health and more people at risk for heart health.

Percent (%) of Treadmill Result by Weight/Height Group

Heart Health Specific ● At risk ● Extremely Healthy ● Healthy ● Normal ● Warning



The above graph was made into four groups using the CASE statement.

Group 1: $[\text{Weight/Height}] < 2.0$ (Overweight)

Group 2: $2.0 \leq [\text{Weight/Height}] \leq 2.49$ (Higher normal)

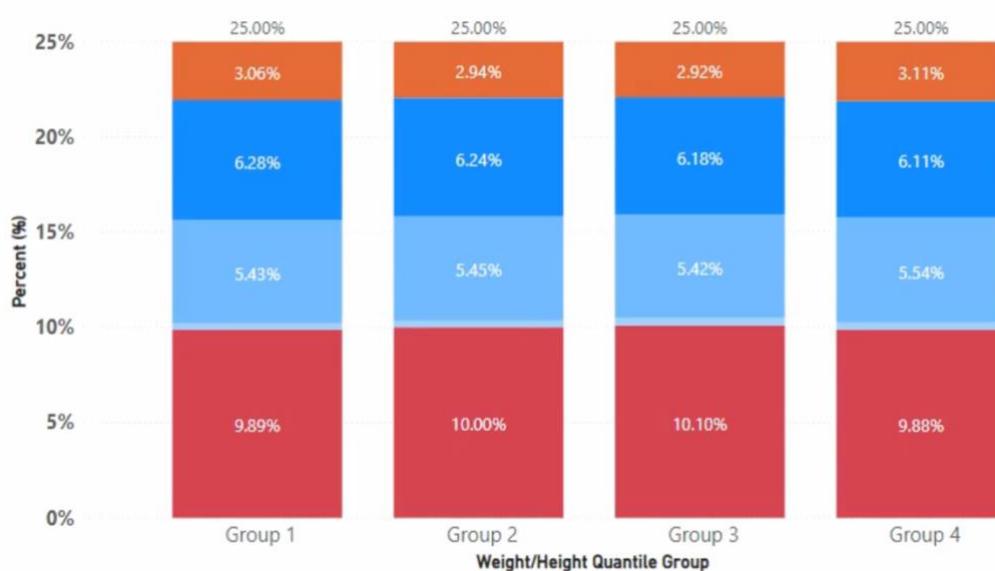
Group 3: $2.50 \leq [\text{Weight/Height}] \leq 2.99$ (Lower normal)

Group 4: $[\text{Weight/Height}] \geq 3.0$ (Thinness)

As a result, the distribution of group 2 is the most distributed at 29.84%, and group 1 is the least distributed at 19.91%. However, it is difficult to compare treadmill results for four groups because the distributions of each group are different.

Percent (%) Treadmill Result by Weight/Height Quantile Group

Heart Health Specific ● At risk ● Extremely Healthy ● Healthy ● Normal ● Warning



The graph above was divided into four groups by height using NTILE() statements to compare with the graph shown earlier (CASE statement graph). It shows that the treadmill results are evenly distributed across each group.

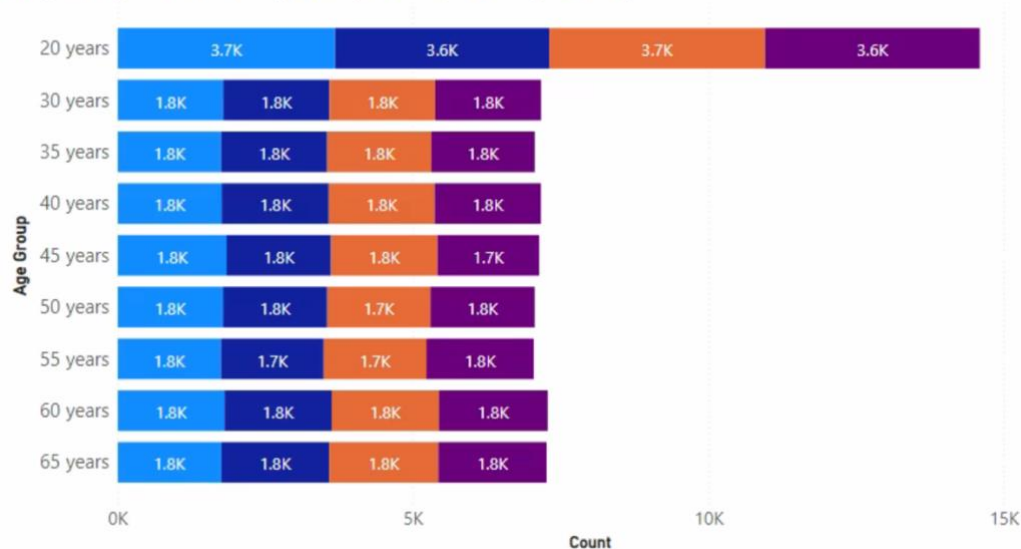
Also, checking the treadmill results for each group, the distribution of heart health in all groups is distributed in the order of at risk > Warning > Normal > Healthy > Extremely Healthy. Also in Group 3, the percentage of people at risk heart health is 10.10%, the normal

rate is 6.18%, and the healthy rate is 5.42%, showing the worst treadmill test. On the other hand, in Group 1, the percentage of people at risk for heart health is 9.89%, the normal rate is 6.28%, and the healthy rate is 5.43%, showing the best results in the treadmill test.

Therefore, comparing these two graphs above, the group created using NTILE statements shows more insightful results than using CASE statements.

Count by Age Group and Weight/Height Quantile Group

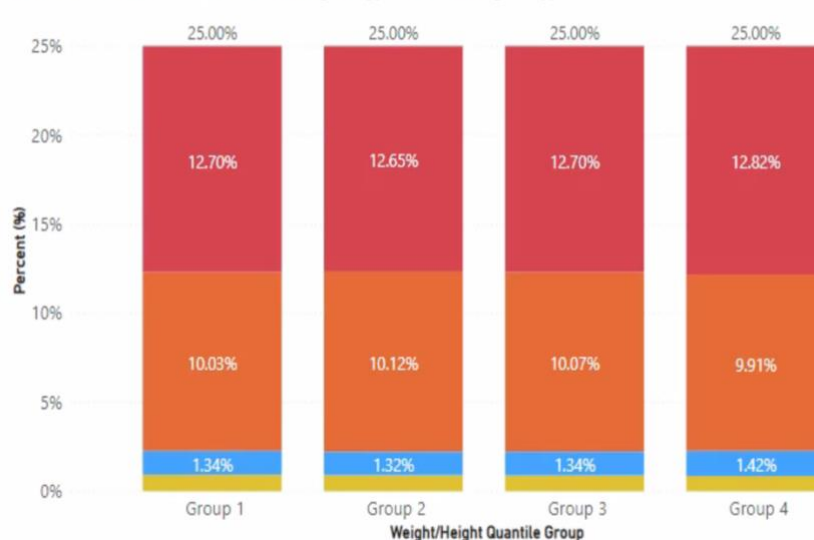
Weight/Height Quantile Group ● Group 1 ● Group 2 ● Group 3 ● Group 4



The above graph shows the distribution of weight/height quantile group by age group, and as a result, the distribution of weight/height groups for all age groups is evenly distributed.

Percent (%) of Blood Pressure by Weight/Height Quantile Group

Blood Pressure ● Elevated ● Normal ● Stage 1 Hypertension ● Stage 2 Hypertension

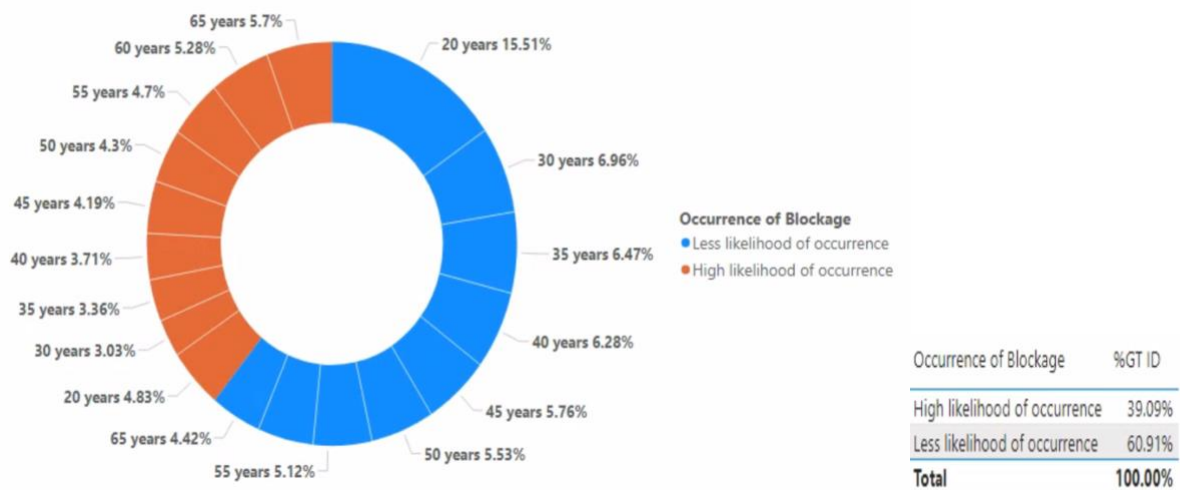


In addition, comparing the blood pressure levels by weight/height quantile group through the graph on the left, state 2 hypertension accounts for more than half of all groups, and people with blood pressure in the normal category show a very small percentage from 1.3% to 1.45% for all groups.

As a result of checking the treadmill result and blood pressure stage for the weight/height group, it was judged that the weight/height group had no significant effect on the treadmill result and blood pressure stage, although there were slight differences between groups.

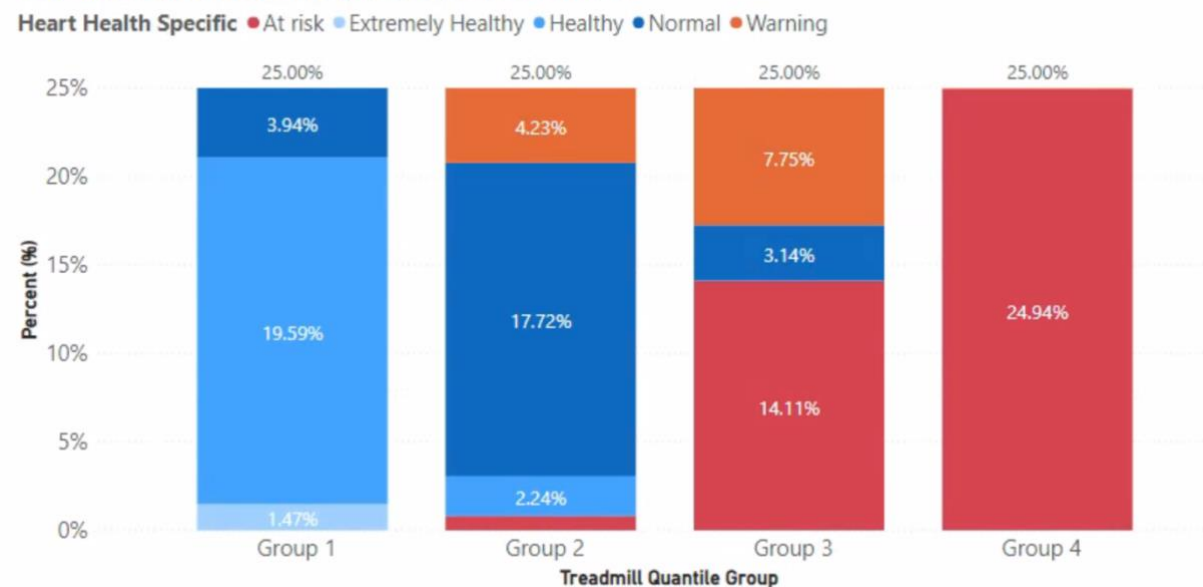
The treadmill test requires walking until the subject's heart rate reaches its peak, and if it exceeds 85% of the average maximum heart rate in the subject's age group, obstruction of the subject's heart may occur. To confirm this, first, check the distribution of blockage that may have occurred while conducting the test in the two groups below.

Percentage by Occurrence of Blockage and Age Group



From the pie chart and metrics above, the rate at which blockage could have occurred in the subject was 39.09%, and the rate at which blockage did not occur in the subject was 60.91%. In addition, because of checking the rate at which blockages have occurred by age, as the age group increases, the rate at which blockages have not occurred decreases and the rate at which blockages have occurred gradually increases.

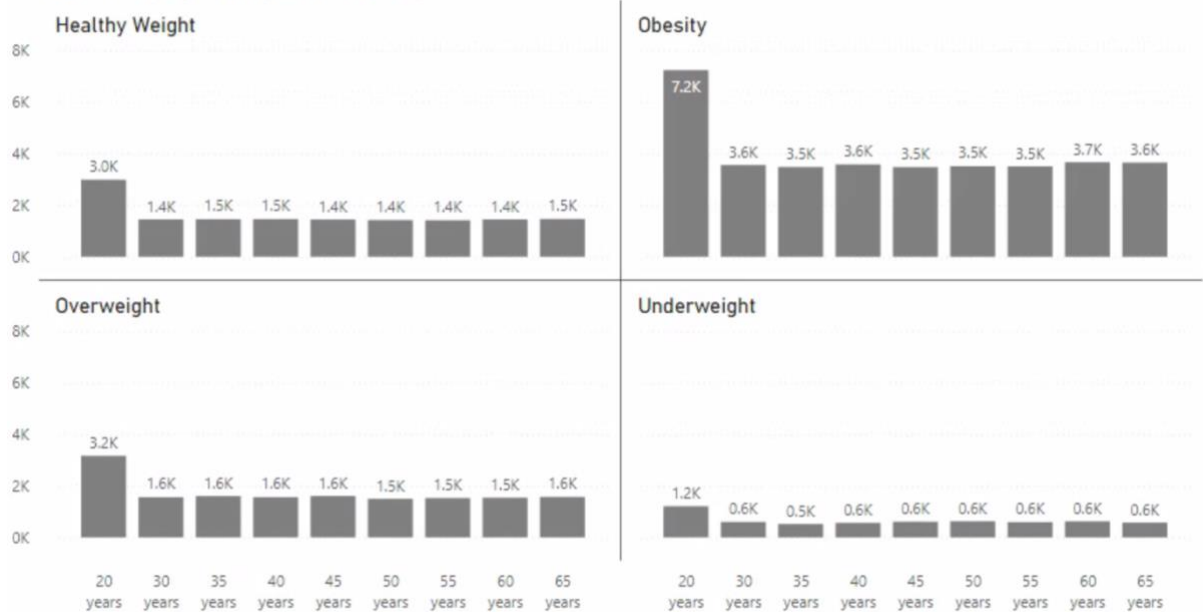
Percent (%) by Heart Health by Treadmill Quantile Group



The above graph was made into four groups using NTILE statements based on the treadmill test, and heart health status was confirmed through this. As a result, 19.59% of the total 25% in Group 1 are in healthy heart conditions, 17.72% of the total 25% in Group 2 are in normal heart conditions, and 24.94% of the total 25% in Group 4 are classified as subjects with at-risk heart conditions.

This part checks how BMI affects the treadmill and blood pressure levels.

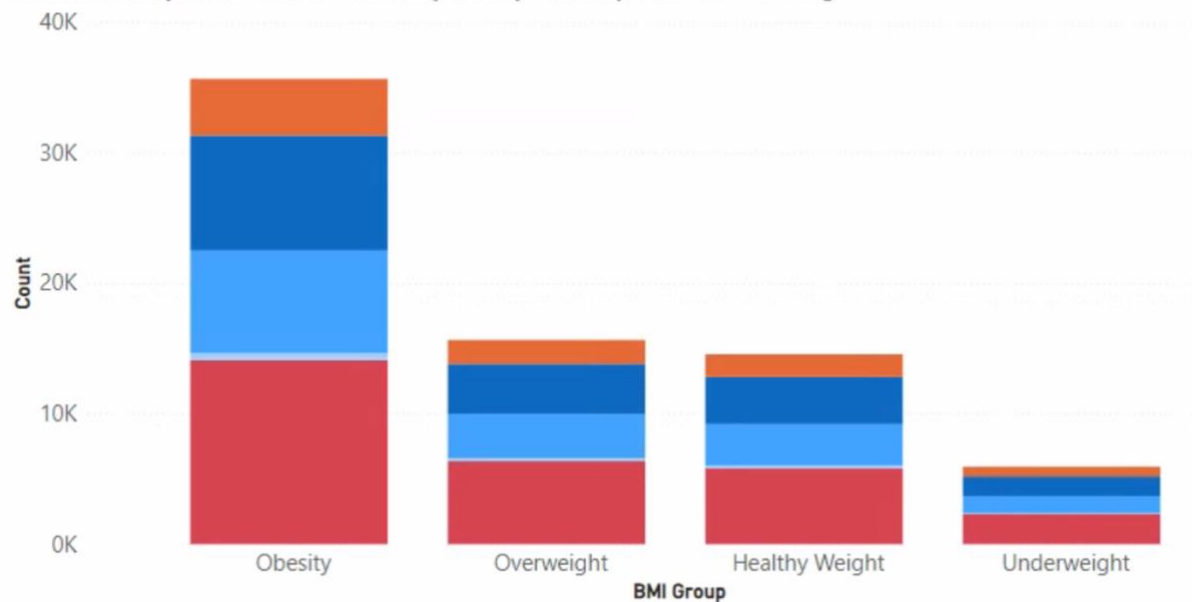
Distribution by Age Group and BMI Group



First, the graph above shows that people in their 20s are the most distributed in all BMI groups that all age groups are the most distributed in the obesity group and all age groups are the least distributed in the underweight group.

Count of Treadmill Result by BMI Group

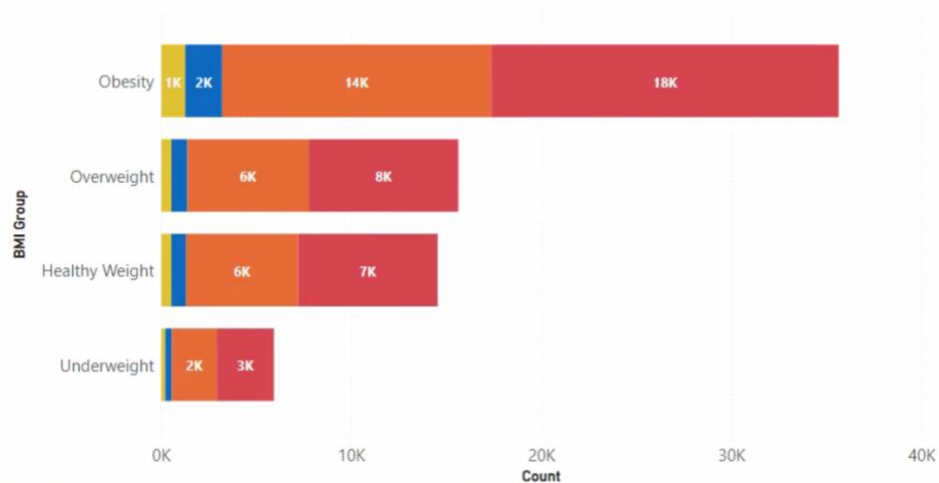
Heart Health Specific ● At risk ● Extremely Healthy ● Healthy ● Normal ● Warning



In addition, because of checking the heart health status by BMI group through the above-stacked chart, it can be confirmed that the heart health status is at risk most in all groups.

Count of Blood Pressure by BMI Group

Blood Pressure ● Elevated ● Normal ● Stage 1 Hypertension ● Stage 2 Hypertension



The above-stacked chart shows blood pressure levels by BMI group. As a result, in all BMI groups, stage 2 of hypertension is the most distributed and the normal range is the least distributed.

Considering that almost all BMI groups are distributed at similar rates in treadmill results and blood pressure, it is judged that BMI and treadmill results are not correlated and also BMI and blood pressure are not correlated.

Heart Health	Average of BMI	Average of Age
At risk	28.88	48.31
Warning	28.93	43.16
Normal	28.83	43.05
Healthy	28.91	40.93
Extremely Healthy	29.31	26.29
Total	28.89	44.45

Finally, the above matrix shows the average BMI and average age for the heart health group (Treadmill test result). As a result, it can be confirmed that the average BMI of all groups of heart health is between 28.93 and 29.31, so it is 'overweight'. If it is associated with BMI and heart health, BMI should increase as heart health deteriorates. But the matrix above shows that the average BMI does not increase, indicating that there is no association between heart health and BMI. And if you look at the average age of each heart health group, you can see that the average age increases as heart health decreases, which shows there is a correlation between heart health and age.

Conclusion

As a result of analyzing the relationship between each patient's health record variables through a treadmill stress test, the sex ratio of recorded patients was almost the same as 49.91% for men and 50.09% for women, and almost all age groups belong to obesity. And more in-depth analysis confirmed that weight/height, BMI, and blood pressure were not correlated with the treadmill test result. On the other hand, the lower the age group, the higher the distribution of groups with healthy hearts, and as the age group increases, the distribution of groups whose heart health is at risk increases. This shows that there is a correlation between age and heart health. (Treadmill test result).

Appendix

Part One

USE [Featherman_Analytics];

-- Start with a query from the prior related assignment3 and load an array using the data in the table shown above.

-- create the HealthHeartExperimental table

```
DECLARE @HealthHeartExperimental TABLE ([ID] INT, [Age] INT, [SexNum] INT,
[SysBP] DECIMAL(8,2), [DiaBP] DECIMAL(8,2), [HRTreadmillTest] DECIMAL(8,2)
, [WeightLbs] DECIMAL(8,2), [HeightInches] DECIMAL(8,2), [BMI] DECIMAL, [Age
Group] NVARCHAR(50), [BMI Group] NVARCHAR(50)
, [AvgMaxHeartRate] DECIMAL(8,2), [HR Zone 50] DECIMAL(8,2), [HR Zone 60]
DECIMAL(8,2), [Heart Health] NVARCHAR(50)
,[Blood Pressure] NVARCHAR(50), [Gender] NVARCHAR(50))
```

-- insert data into the table (copy query from assignment 3)

```
INSERT INTO @HealthHeartExperimental ([ID], [Age], [SexNum], [SysBP], [DiaBP],
[HRTreadmillTest]
```

```
, [WeightLbs], [HeightInches], [BMI], [Age Group], [BMI GROUP]
```

```
, [AvgMaxHeartRate], [HR Zone 50], [HR Zone 60], [Heart Health], [Blood Pressure],
[Gender])
```

```
(SELECT *
```

```
, CASE -- results of treadmill test
```

```
    WHEN [HRTreadmillTest] >= [AvgMaxHeartRate] THEN 'At risk'
```

```
    WHEN [HRTreadmillTest] < [AvgMaxHeartRate] AND
```

```
    [HRTreadmillTest] > [HR Zone 60]
```

```
    THEN 'Healthy'
```

```
    WHEN [HRTreadmillTest] <= [HR Zone 50] OR [HRTreadmillTest] <= [HR Zone
```

```
60]
```

```
    THEN 'Extremely Healthy'
```

```
    END AS [Heart Health]
```

```
, CASE -- Blood Pressure
```

```
    WHEN [SysBP] < 120 AND [DiaBP] < 80 THEN 'Normal'
```

```
    WHEN ([SysBP] BETWEEN 120 AND 129) AND [DiaBP] < 80 THEN 'Elevated'
```

```
    WHEN ([SysBP] BETWEEN 130 AND 139) OR ([DiaBP] BETWEEN 80 AND 89)
```

```
THEN 'Stage 1 Hypertension'
```

```
    WHEN ([SysBP] >= 140 AND [SysBP] < 180) OR ([DiaBP] >= 90 AND [DiaBP] <
```

```
120) THEN 'Stage 2 Hypertension'
```

```
    WHEN [SysBP] >= 180 OR [DiaBP] >= 120 THEN 'Hypertensive crisis'
```

```
    END AS [Blood Pressure]
```

```
, CASE -- Gender
```

```
    WHEN sex = 1 THEN 'Male'
```

```
    ELSE 'Female'
```

```
    END AS [Gender]
```

```
FROM
```

```
(SELECT *
```

```
, CASE -- BMI groups
```

```
    WHEN [BMI] < 18.5 THEN 'Underweight'
```



```

        WHEN [BMI] BETWEEN 18.5 AND 24.9 THEN 'Healthy Weight'
        WHEN [BMI] BETWEEN 25.0 AND 29.9 THEN 'Overweight'
        WHEN [BMI] >= 30.0 THEN 'Obesity'
    END AS [BMI Group]
, CASE -- Avg Max Heart Rate (85%)
    WHEN [Age Group] = '20 years' THEN 170
    WHEN [Age Group] = '30 years' THEN 162
    WHEN [Age Group] = '35 years' THEN 157
    WHEN [Age Group] = '40 years' THEN 153
    WHEN [Age Group] = '45 years' THEN 149
    WHEN [Age Group] = '50 years' THEN 145
    WHEN [Age Group] = '55 years' THEN 140
    WHEN [Age Group] = '60 years' THEN 136
    WHEN [Age Group] = '65 years' THEN 132
    WHEN [Age Group] = '70 years' THEN 128
    END AS [AvgMaxHeartRate]
, CASE -- HR Zone 50%
    WHEN [Age Group] = '20 years' THEN 100
    WHEN [Age Group] = '30 years' THEN 95
    WHEN [Age Group] = '35 years' THEN 93
    WHEN [Age Group] = '40 years' THEN 90
    WHEN [Age Group] = '45 years' THEN 88
    WHEN [Age Group] = '50 years' THEN 85
    WHEN [Age Group] = '55 years' THEN 83
    WHEN [Age Group] = '60 years' THEN 80
    WHEN [Age Group] = '65 years' THEN 78
    WHEN [Age Group] = '70 years' THEN 75
    END AS [HR Zone 50]
, CASE -- HR Zone 60%
    WHEN [Age Group] = '20 years' THEN 120
    WHEN [Age Group] = '30 years' THEN 114
    WHEN [Age Group] = '35 years' THEN 111
    WHEN [Age Group] = '40 years' THEN 108
    WHEN [Age Group] = '45 years' THEN 105
    WHEN [Age Group] = '50 years' THEN 102
    WHEN [Age Group] = '55 years' THEN 99
    WHEN [Age Group] = '60 years' THEN 96
    WHEN [Age Group] = '65 years' THEN 93
    WHEN [Age Group] = '70 years' THEN 90
    END AS [HR Zone 60]
FROM
(SELECT [ID],[age],[sex],[SysBP],[DiaBP],[HRTreadmillTest],[weightLbs]
,[heightInches],[BMI]
, CASE -- age groups
    WHEN age BETWEEN 20 AND 29 THEN '20 years'
    WHEN age BETWEEN 30 AND 34 THEN '30 years'
    WHEN age BETWEEN 35 AND 39 THEN '35 years'
    WHEN age BETWEEN 40 AND 44 THEN '40 years'
    WHEN age BETWEEN 45 AND 49 THEN '45 years'
    WHEN age BETWEEN 50 AND 54 THEN '50 years'

```

```

        WHEN age BETWEEN 55 AND 59 THEN '55 years'
        WHEN age BETWEEN 60 AND 64 THEN '60 years'
        WHEN age BETWEEN 65 AND 69 THEN '65 years'
        WHEN age BETWEEN 70 AND 79 THEN '70 years'
    END AS [Age Group]
FROM [featherman].[Health_heart_experimental]
) AS data
) AS data2
)

-- 1. create the table (ReachedMaxHeartRate)
-- this table based on their actual max heart rate on the treadmill test as compared to the max
their age group.
DECLARE @ReachedMaxHeartRate TABLE ([ID] INT, [Heart Health Specific]
NVARCHAR(50))

-- insert the data into the table
INSERT INTO @ReachedMaxHeartRate ([ID], [Heart Health Specific])
SELECT [ID]
, CASE -- reached or exceeded max heartrate
    WHEN [HRTreadmillTest] >= [AvgMaxHeartRate] THEN 'At risk'
    WHEN [HRTreadmillTest] >= ([AvgMaxHeartRate] * 0.9) THEN 'Warning'
    WHEN [HRTreadmillTest] >= ([AvgMaxHeartRate] * 0.7) THEN 'Normal'
    WHEN [HRTreadmillTest] >= ([AvgMaxHeartRate] * 0.5) THEN 'Healthy'
    ELSE 'Extremely Healthy'
END
FROM @HealthHeartExperimental

-- 2.create the table (WeightHeight)
-- this table based on weight pounds divided by height inches
DECLARE @WeightHeight TABLE ([ID] INT, [Weight/Height] DECIMAL(8,2)
, [Weight/Height Group] NVARCHAR(50)
, [Weight/Height Quantile #] DECIMAL, [Weight/Height Quantile Group] NVARCHAR(50))

-- insert the data into the table
INSERT INTO @WeightHeight ([ID], [Weight/Height], [Weight/Height Quantile #],
[Weight/Height Quantile Group])
SELECT [ID], [WeightLbs]/[HeightInches]
, NTILE(4) OVER(ORDER BY ([WeightLbs]/[HeightInches])) -- quartiles based on weight
divided by height
, CASE NTILE(4) OVER(ORDER BY ([WeightLbs]/[HeightInches])) --quartiles lables
    WHEN 1 THEN 'Group 1'
    WHEN 2 THEN 'Group 2'
    WHEN 3 THEN 'Group 3'
    WHEN 4 THEN 'Group 4'
END
FROM @HealthHeartExperimental

--Update categories based on the weight/height
UPDATE @WeightHeight SET [Weight/Height Group] =

```

```

(
CASE
    WHEN [Weight/Height] < 2.0 THEN 'Group 1'
    WHEN [Weight/Height] BETWEEN 2.0 AND 2.49 THEN 'Group 2'
    WHEN [Weight/Height] BETWEEN 2.50 AND 2.99 THEN 'Group 3'
    WHEN [Weight/Height] >= 3.0 THEN 'Group 4'
END
)

-- 3.create the table (Blockage)
-- this table based on treadmill results
DECLARE @Blockage TABLE ([ID] INT
, [Occurrence of Blockage] NVARCHAR(50)
, [Treadmill Quantile #] DECIMAL, [Treadmill Quantile Group] NVARCHAR(50))

-- insert the data into the table
INSERT INTO @Blockage ([ID], [Occurrence of Blockage], [Treadmill Quantile #],
[Treadmill Quantile Group])
SELECT [ID]
, CASE -- categories about occurrence of blockage
    WHEN [HRTreadmillTest] > [AvgMaxHeartRate] THEN 'High likelihood of
occurrence'
    WHEN [HRTreadmillTest] <= [AvgMaxHeartRate] THEN 'Less likelihood of
occurrence'
END
, NTILE(4) OVER(ORDER BY ([HRTreadmillTest])) -- quartiles based on treadmill test
, CASE NTILE(4) OVER(ORDER BY ([HRTreadmillTest])) -- quartiles labels
    WHEN 1 THEN 'Group 1'
    WHEN 2 THEN 'Group 2'
    WHEN 3 THEN 'Group 3'
    WHEN 4 THEN 'Group 4'
END
FROM @HealthHeartExperimental

--merge all arrays by ID key and show the results
SELECT hhe.[ID], [Age], [SysBP], [DiaBP], [HRTreadmillTest], [WeightLbs],
[HeightInches], [BMI], [Age Group], [BMI Group]
, [AvgMaxHeartRate], [Blood Pressure], [Gender], [Heart Health Specific], [Weight/Height]
, [Weight/Height Group], [Weight/Height Quantile #],[Weight/Height Quantile Group]
, [Occurrence of Blockage], [Treadmill Quantile #], [Treadmill Quantile Group]
FROM @HealthHeartExperimental as hhe
INNER JOIN @ReachedMaxHeartRate as rmr
ON hhe.ID = rmr.ID
INNER JOIN @WeightHeight as wh
ON hhe.ID = wh.ID
INNER JOIN @Blockage as bk
ON hhe.ID = bk.ID

```