Title Page

Contents

# Background and Introduction

### Aim

The aim of this report is to first conduct an analysis of health score data to ascertain the relative health scores of a population. Then to use the results of the analysis to drive predictive models that will help identify the health scores of individuals. The data used for analytics covers:

* Age
* Sex
* Weight
* Height
* IQ
* Units of Alcohol Consumed per Day
* Cigarettes Consumed per Day
* Level of Activity
* Final Health Score

### Content

The analytical requirements have been set in the form of three questions:

* Are there any significant differences between different segments of the population in terms of their lifestyle choices (for example, male and female, different age groups)?
* Which individual characteristics and lifestyle choices impact a person’s health score (and to what extent)?
* What would be the impact on the overall population (in terms of health score) if nobody consumed alcohol and did not smoke?

The predictive requirements have been set as predicting the health scores of 20 individuals based on the health data supplied.

### Conclusions

# Analysis

### Analytical Techniques Chosen

Techniques will be aimed at answering the three questions posed in turn:

* Comparison of health score with each data value available (alcohol consumption, male/female)
* In generalised terms for age groups and other highly variable factors with a look into the depth of impact each factor has
* A look at the impact on the population if there was no smoking and drinking in a population

### Results of Analytical Techniques

#### Age

The ages within the data vary from between 18 and 80, obtained using min and max functions. I have separated the ages into the age groups of 18-30, 31-40, 41-50, 51-60, 61-70, 71-80.

The mean health score of each age group was as follows:

* 18-30: 238.669912366
* 31-40: 196.799758745
* 41-50: 167.826923077
* 51-60: 148.592405063
* 61-70: 129.027918782
* 71-80: 116.848407643

The 18-30 age group had a 204.3% higher mean value than the 71-80 age group.

The median health score of each age group was as follows:

* 18-30: 225.0
* 31-40: 188.0
* 41-50: 159.0
* 51-60: 143.5
* 61-70: 121.0
* 71-80: 109.0

The 18-30 age group had a 206.4% higher median value than the 71-80 age group.

#### Male/Female

The mean health score of males and females were:

* Males: 141.353159851
* Females: 187.351091142

Females are therefore 32.5% more healthy on average than males.

The median results were:

* Males: 131.0
* Females: 174.0

Females having a 32.8% higher median value than males.

#### Weight

The weight within the data varies from between 77 pounds and 217 pounds. Division into 5 equal groups gives a 28-pound range for each group.

The mean health score of each weight group was as follows:

* 77 - 104: 201.438016529
* 105 - 132: 190.261529809
* 133 - 160: 167.562848297
* 161 - 188: 141.699077955
* 189 - 217: 136.023391813

The 77 – 104 weight group had a 148.1% higher mean value than the 189 – 217 weight group.

The median health score of each weight group was as follows:

* 77 - 104: 188.0
* 105 - 132: 177.0
* 133 - 160: 155.0
* 161 - 188: 131.0
* 189 - 217: 130.0

The 77 – 104 weight group had a 144.6% higher median value than the 189 – 217 weight group.

#### Height

The height within the data varies from between 56 inches and 83 inches. I’ve divided this into 4 groups of 7 inches with the 1st group containing 6.

The mean health score of each height group was as follows:

* 56 - 62: 173.597938144
* 63 - 69: 181.708484409
* 70 - 76: 152.313771063
* 76 - 83: 146.318181818

The 63 – 69 height group had a 124.2% higher mean value than the 76 – 83 height group. This is the first result that does not have an upper and lower bound on the extremes.

The median health score of each height group was as follows:

* 56 - 62: 161.0
* 63 - 69: 167.0
* 70 - 76: 138.0
* 76 - 83: 135.5

The 56 – 62 height group had a 123.2% higher median value than the 76 – 83 height group.

#### IQ

The IQ within the data varies from between 79 and 122. Division into 4 groups gives an 11-point range for each group except for the 1st group which has 10 points.

The mean health score of each IQ group was as follows:

* 79 - 89: 168.286486486
* 90 - 100: 172.073260073
* 101 - 111: 168.031151242
* 112 - 122: 167.943661972

The 90 – 100 IQ group had a 102.4% higher mean value than the 112 – 122 IQ group. Another case where it is not the extremes that have the highest contrast.

The median health score of each IQ group was as follows:

* 79 - 89: 155.0
* 90 - 100: 159.0
* 101 - 111: 154.0
* 112 - 122: 153.0

The 90 – 100 IQ group had a 103.9% higher median value than the 112 – 122 IQ group.

#### Units of Alcohol Consumed Per Day

#### Cigarettes Smoked Per Day

#### Level of Activity

# Predictions

### Predictive Techniques Chosen

### Results of Predictive Techniques

# Findings and Recommendations

# Appendix

### Code Listing