Phase 2

1) "Will financial support reduce obesity?"

2) Domain: Health and Finance

3) Consumptions of food to compensate for stress is called stress eating and is a term commonly used nowadays. In this study, we are aiming to find the relationship between financial stress and obesity due to stress eating. The government, health sector and finance sector will be interested in this relation since obesity is the main reason for many serious health conditions. Reducing the obesity population will be an excellent accomplishment itself, moreover it will also decrease the expenses at hospitals, thus benefiting the government as well.

4) There aren't any reports regarding the relationship between financial stress and obesity in Victoria. We aim to show this relationship in visual and quantitative formats using reliable data. Furthermore, we will divide the population into male and female and see how financial stress affects each gender. We plan on studying the behavior of the percentage of obese population when there's no financial stress, so that we could check if any of the solutions we derive are fruitful.

5) The following Datasets are used:

- 2011 LGA Health Risk Factors (Mar/2011-July/2013): Modelled estimate of several health risk factors by LGA. This Dataset contains the estimated data of the percentage and count of persons who are experiencing health risk factors. It comes with a reliability rating for each estimation. Dataset downloaded from Aurin. https://portal.aurin.org.au/("personal_financial_stressors.csv").

- 2011 LGA Personal and Financial Stressors (Nov/2010). This dataset contains the percentage and count of persons who are having financial problems by LGA (e.g. Cash flow problems). It comes with a statistical significance(reliability) rating for each estimation. Dataset downloaded from Aurin. https://portal.aurin.org.au/("health_risk_factors_estimates.csv")

The 2 datasets are from different years because stress cannot cause obesity immediately, it may take months or years for a person's mass to be in the obese zone. In this case, the financial stress caused in 2010 would affect the number of people who are obese in 2011,2012 and 2013.

The rows with "Unincorporated Vic" as the LGA name was removed since 2011 LGA Personal and Financial Stressors csv file did not contain data for "Unincorporated Vic" LGA. Therefore, having only 79 LGAs for comparison is a limitation.

7) The following 2 operations were done using Microsoft Excel:

1. 2011 LGA Health Risk Factors csv file contains data relating to all the health risk factors therefore, we created a new csv file named "new_health_risk.csv" with the

columns only pertaining to obesity. Furthermore, we renamed the columns with shorter names with '_' instead of spaces.

2. 2011 LGA Personal and Financial Stressors csv file contains data with many different forms of Personal and Financial and personal stressors. We therefore, created a csv file with the columns we need called "new_stress.csv". This csv file contains columns count, percentage and the statistical significance of:

   o Cash flow problem in the last 12 months.

   o Could Raise 2000 dollars within a week.

   o Government support as main source of income in the last 2 years

   o At least one dissaving action in the last 6 months.

   Plus, columns were renamed with shorter names with '_' instead of spaces.

Afterwards, we used the merge function in python to inner join the two csv files according to the LGA code. We created a scatter plot of the percentages of people who faced cash flow problems in the last 12 months vs the percentages of people who are obese by LGA.
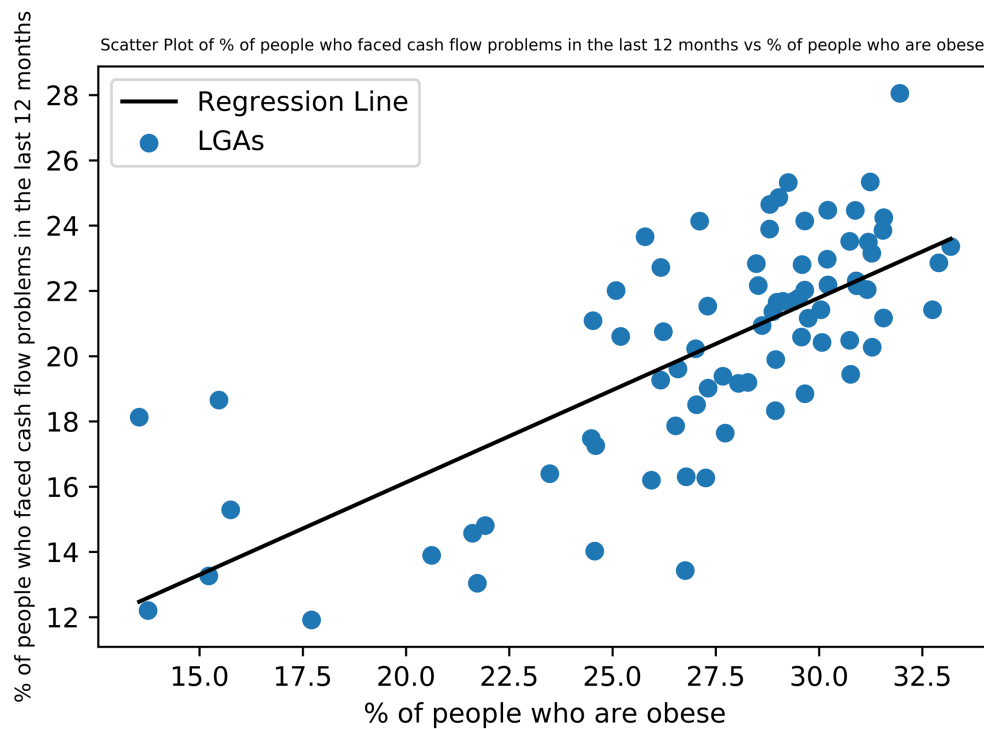

FIGURE 1

The dots represent each LGA and the line represents the line of best fit (Line of Regression). The line of regression and Pearson correlation coefficient was calculated using the percentages. The scatter plot clearly shows us that whenever the percentage of people who faced cash flow problems in the last 12 months increased, the percentage of obese population increased as well. To back up this observation, we found that the Pearson correlation coefficient is 0.717.

The following table (TABLE 1) shows what the Pearson correlation Coefficient means:

Coefficient	Interpretation
-1	Perfect Negative linear relationship
-0.70	Strong Negative linear relationship
-0.50	Moderate Negative linear relationship
-0.30	Weak Negative linear relationship
0	No linear relationship
+0.30	Weak positive linear relationship
+0.50	Moderate positive linear relationship
+0.70	Strong positive linear relationship
+1	Perfect positive linear relationship

TABLE 1

According to this table, the coefficient we calculated falls in between strong and perfect positive linear relationship. It is impossible to receive a coefficient of 1 or -1 in real life due to unexpected events (e.g. genetic obesity). Therefore, coefficient 0.717 shows an extremely good linear relationship between financial stress and obesity.

Furthermore, we drew the line of best fit (regression line) to show the linear relationship clearly. To make sure the data we processed were reliable, we calculated the average reliability rating of the column for both the columns used in the above calculation.

Column Name	Average Reliability Rating
Percentage of people who faced cash flow problems in the last 12 months	2.95
Percentage of people who are obese	3.0

TABLE 2

Reliability rating may vary from 1 to 3, where 1 being non-reliable and 3 being reliable. Since both the average reliability rating are 3 or extremely close to 3, we could conclude that the data used is highly reliable.

We have used percentages for the correlation rather than the number of people for a specific reason. The number of people with a specific criterion for a certain LGA will depend on the total population size of the LGA. Therefore, if we had checked for the correlation of the 2 population counts it would surly correlate even if the 2 criteria didn't correlate. So, by using percentages, we normalize the different LGA population sizes and make a fair comparison.

8) As you can see, the data we have used is reliable and the values used are unbiased by population sizes, therefore the results obtained are well grounded. The average percentage of population classified as obese in all the LGAs in Victoria is 27.35%, that is more than a quarter of the population. Obesity leads to many severe and deadly illnesses, treating these illnesses costs the government financially. Finding a way to decrease the percentage of population who are obese will benefit citizens and improve the mortality rate. Furthermore, it will reduce hospital expenses, thus benefiting the government financially. So far, the results look promising, only further analysis will reveal whether "financial stress reduces obesity?"