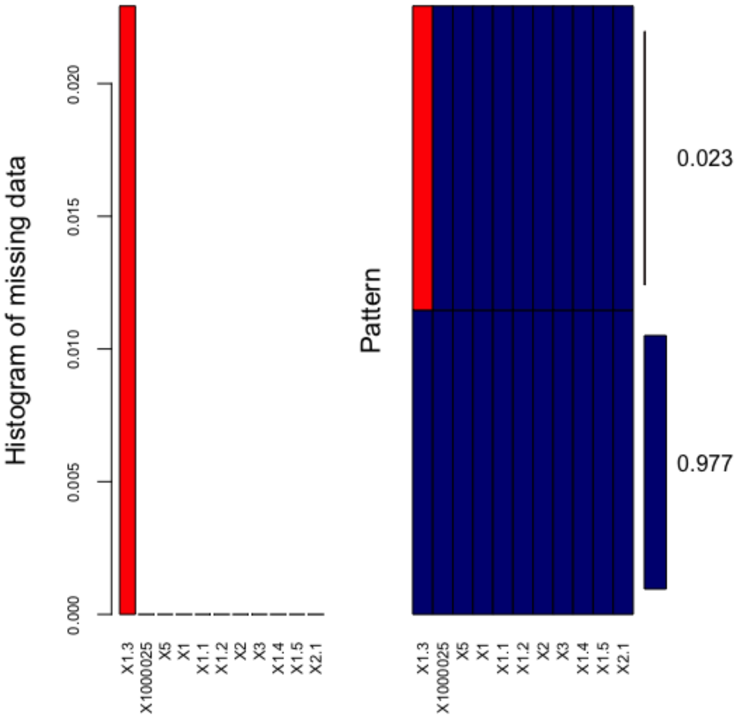
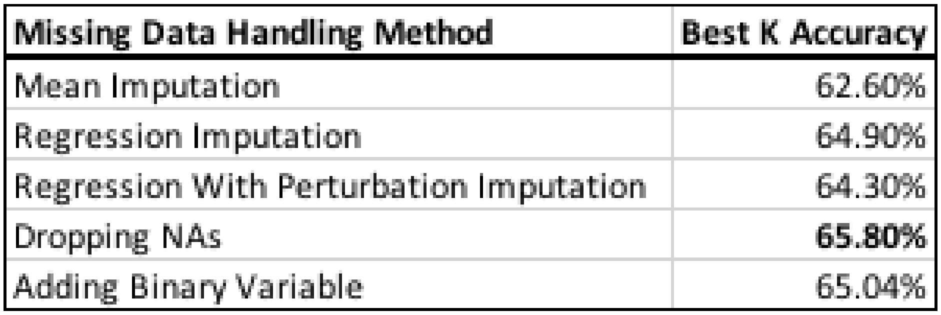
14.1 – Please review the included R file 14.1.R along with this writeup

First I loaded and inspected the data, and saw that column 7 had ? for some values. I replaced the ? with NAs to make it easier to impute and replace the data later. I then created a histogram of the missing data to see what was missing. As the plot below shows, only column 7/X1.3 is missing data, and it's only a small amount (about 2%) of the total data.



I used the mice package to impute the missing values using mean, regression, and regression with perturbation and created separate datasets using each of those methods (lines 25-30 of the attached R code). I also created a dataset that just omitted rows containing NAs and a data set with a binary variable added that indicates missing values.

I then trained KNN models on each of the datasets and used cross validation to see which model was the highest quality.



Missing Data Handling Method	Best K Accuracy
Mean Imputation	62.60%
Regression Imputation	64.90%
Regression With Perturbation Imputation	64.30%
Dropping NAs	**65.80%**
Adding Binary Variable	65.04%

Interestingly, the best result was from simply removing the rows that had NA values, although all of the results are relatively close.

15.1
I am currently in the process of specifying the components of a new racing bike and would like to make it as light and aerodynamic as possible while staying below my budget and fitting my body dimensions. I would need the weight, coefficient of drag, size and cost for each option. The objective function would be to minimize cost while selecting each of the required parts. The major constraints would be minimizing weight and drag. Other minor constraints that might be worth adding are sizing constraints on certain parts to ensure that they would fit within a certain range that I know is appropriate for me, otherwise the objective function might choose the smallest part possible to optimize for weight.

15.2

For 15.2.1 and 15.2.2 the included Jupyter notebooks have commented code that walks through the process and shows all functions and constraints.

For 15.1 my result was:

```
Foods_Celery,_Raw = 52.64371
Foods_Frozen_Broccoli = 0.25960653
Foods_Lettuce,Iceberg,Raw = 63.988506
Foods_Oranges = 2.2929389
Foods_Poached_Eggs = 0.14184397
Foods_Popcorn,Air_Popped = 13.869322
```

For 15.2, I added the following constraints:

Min and Max for chosen foods:

```
for f in foods:
diet += foodvars[f] >= .1*chosenvars[f]
for f in foods:
diet += foodvars[f] <= 999999999*chosenvars[f]
```

Only pick one of Brocolli and Celery:

```
diet += chosenvars['Frozen Broccoli'] + chosenvars['Celery, Raw'] <= 1, 'Broc/Celery R
```

Require at least 3 or meat/fish/poultry/eggs:

```
diet += chosenvars['Roasted Chicken'] + chosenvars['White Tuna in Water']  +
chosenvars['Poached Eggs'] + chosenvars['Scrambled Eggs'] + chosenvars['Bologna,Turkey] +
chosenvars['Frankfurter, Beef'] + chosenvars['Ham,Sliced,Extralean']  +
chosenvars['Kielbasa,Prk'] + chosenvars['Hamburger W/Toppings']  + chosenvars['Hotdog,
Plain'] + chosenvars['Taco'] + chosenvars['Pork']  + chosenvars['White Tuna in Water'] +
chosenvars['Sardines in Oil']  + chosenvars['Chicknoodl Soup'] + chosenvars['Splt
Pea&Hamsoup']  + chosenvars['Vegetbeef Soup'] + chosenvars['Neweng Clamchwd']  +
chosenvars['Beanbacn Soup,W/Watr'] >= 3, 'Protein Req'
```

Which changed the result to:

```
Foods_Celery,_Raw = 42.399358
Foods_Kielbasa,Prk = 0.1
Foods_Lettuce,Iceberg,Raw = 82.802586
Foods_Oranges = 3.0771841
Foods_Peanut_Butter = 1.9429716
Foods_Poached_Eggs = 0.1
Foods_Popcorn,Air_Popped = 13.223294
Foods_Scrambled_Eggs = 0.1
```