

## Assignment 2

Josh Barber

barb4630@mylaurier.ca

215854630

1. How can we insert a new row of values before the first row of DataFrame object we load from crx\_data.csv in Section 3? (3 points)

There are a couple of ways to accomplish this. A common way is to create two dataframes; one dataframe which only contains the row you want to insert at the top, let us call this df1. Then the other dataframe which will contain the rest of the data, let us call this df2. You can then concatenate both dataframes together, with the pandas concat function. In this case you want to concatenate df1 to df2. Such as `pd.concat([df1, df2]).reset_index(drop=True)`

I however chose to use `loc[-1]` which indexes before the first row in the dataframe. Then I added the values to this row. I then incremented all the other indexes including the new inserted row, then sorted the data in place.

```
import pandas as pd
import os

df = pd.read_csv(os.path.join(os.path.abspath(''), 'crx_data.csv'), header=None)

pd.set_option("display.min_rows", df.shape[0]+1)
pd.set_option('display.max_rows', df.shape[0]+1)

df.loc[-1] = ["b", "24.53", "2.345", "u", "g", "w", "h", "2.75", "t", "t", "4", "t", "s", "00279", "456", "+"]
df.index = df.index + 1
df.sort_index(inplace=True)

print(df.head(5))
print(df.describe())
```

✓ 0.0s

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	b	24.53	2.345	u	g	w	h	2.75	t	t	4	t	s	00279	456	+
1	b	30.83	0.0	u	g	w	v	1.25	t	t	1	f	g	00202	0	+
2	a	58.67	4.46	u	g	q	h	3.04	t	t	6	f	g	00043	560	+
3	a	24.50	0.5	u	g	q	h	1.5	t	f	0	f	g	00280	824	+
4	b	27.83	1.54	u	g	w	v	3.75	t	t	5	t	g	00100	3	+
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
count	691	691	691.0	691	691	691	691	691	691.0	691	691	691	691	691	691	691
unique	3	351	216.0	4	4	15	10	133.0	2	2	24	2	3			
top	b	?	1.5	u	g	c	v	0.0	t	f	0	f	g			
freq	469	12	21.0	520	520	137	399	70.0	362	395	395	374	625			

  

	13	14	15
count	691	691	691
unique	172	241	2
top	00000	0	-
freq	132	295	383