Assignment 6

Group Members:
HAITAO ZHOU (HAZ59)
YINGZHI YANG(YIY50)
XIN JIN(XIJ21)

DATA & METHODS

In this assignment, data is from wikipedia and statista website, which is the same as Franz's data. The website of our data can be found in *References*. We use R Studio to statistic and analyze the data.

The following picture is the screen shot of the data:

id	nation	consumption of tea (kg/yr/capita)	Nobel (/10 million)	chocolate consumption (kg/yr/capita)	alcohol consumption (litre/yr/capita	cigarette consumption (/yr/capita)
1	Australia	0.54	5.006	4.8	12.2	1034
2	Austria	0.21	24.577	7.5	10.3	1650
3	Belgium	0.2	8.85	5.3	11	1455
4	Brazil	1.99	0.048	1.5	8.7	504
5	Canada	0.49	6.4	3.5	10.2	809
6	China	0.92	0.065	0.1	6.7	1711
7	Denmark	0.3	24.695	6.7	11.4	1413
8	Finland	0.31	7.268	5.5	12.3	671
9	France	0.27	9.473	4.2	12.2	854
10	Germany	0.54	13.013	7.5	11.8	1045
11	Greece	0.4	1.826	3.9	10.3	2996
12	Ireland	1.88	14.93	9.6	11.9	1006
13	Italy	0.1	3.345	2.7	6.7	1475
14	Japan	0.97	1.896	1	7.2	1841
15	Netherlands	0.98	11.226	5.2	9.9	801
16	Norway	0.29	24.947	6.9	7.7	534
17	Poland	0.92	3.108	4.7	12.5	1586
18	Portugal	0.11	1.932	1.5	12.9	1114
19	Spain	0.11	1.735	2.1	11.2	1757
20	Sweden	0.41	30.677	5.6	9.2	715
21	Switzerland	0.3	30.125	9.5	10.7	1722
22	UK	2.06	19.315	11	11.6	750
23	USA	0.53	10.97	4.6	9.2	1028

Analysis

1.Summary of Data:

First of all, we show some summary information about 'Nobel.laureates', 'consumption of tea', 'chocolate consumption', 'alcohol' and 'cigarette'.

Consumption.of.tea.kg.yr.capita.
Min.   :0.1000
1st Qu.:0.2800
Median :0.4100
Mean   :0.6448
3rd Qu.:0.9200
Max.   :2.0600

Nobel.laureates /10.million
 Min.   : 0.048
 1st Qu.: 2.520

```
 Median : 8.850
 Mean   :11.106
 3rd Qu.:17.122
 Max.   :30.677


Chocolate.consumption kg/yr/capita.
Min.   : 0.100
1st Qu.: 3.100
Median : 4.800
Mean   : 4.996
3rd Qu.: 6.800
Max.   :11.000

alcohol.consumption...litre.yr.capita.
 Min.   : 6.70
 1st Qu.: 9.20
 Median :10.70
 Mean   :10.34
 3rd Qu.:11.85
 Max.   :12.90

cigarette.consumption....yr.capita.
 Min.   : 504
 1st Qu.: 805
 Median :1045
 Mean   :1238
 3rd Qu.:1618
 Max.   :2996
```
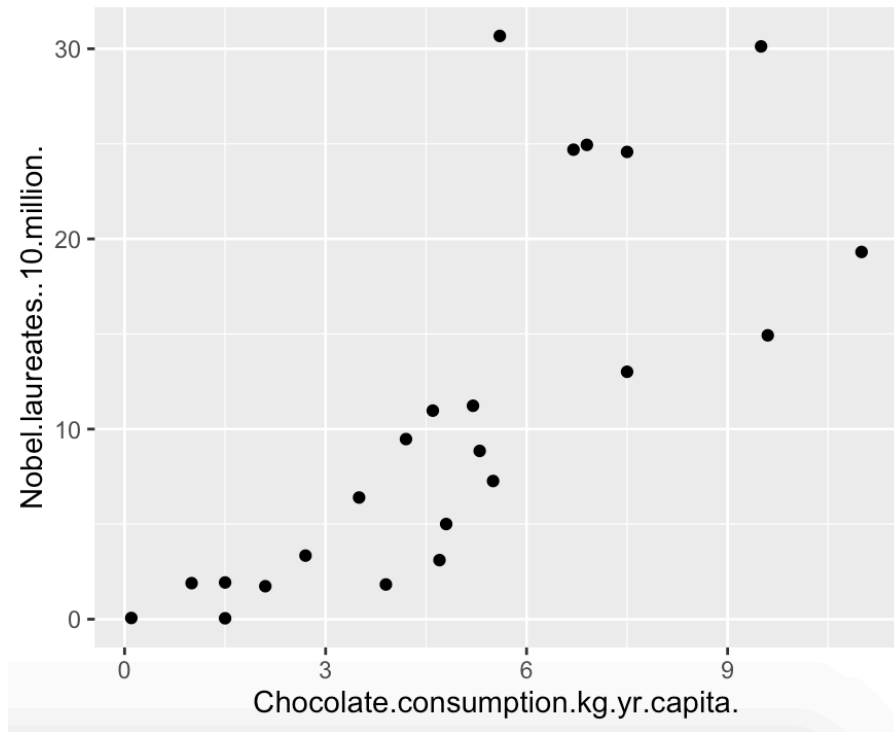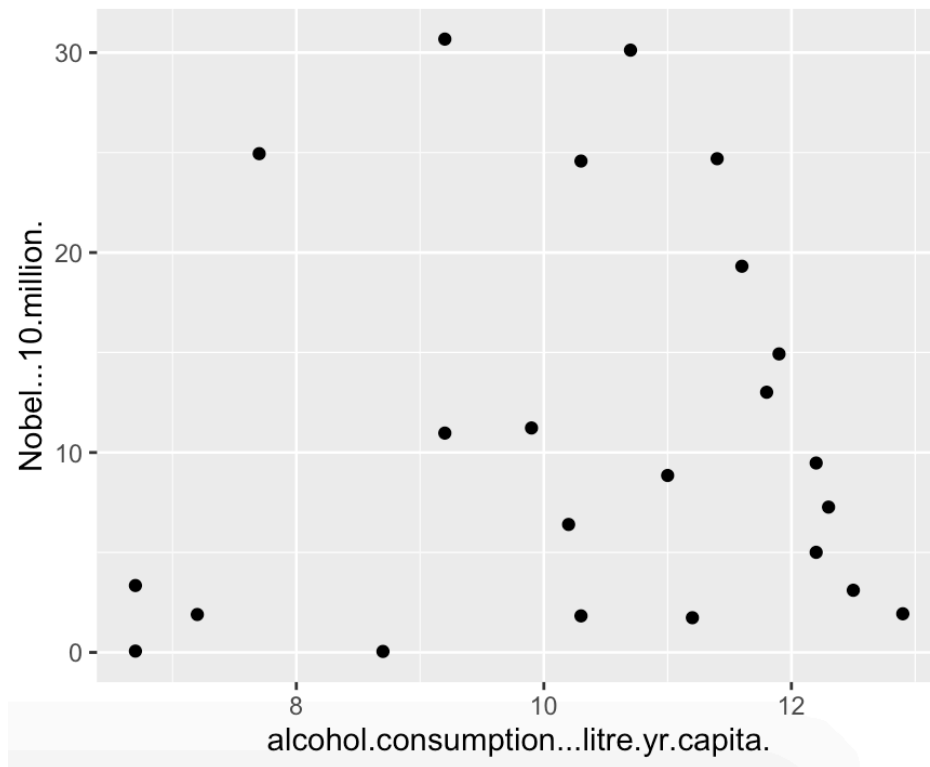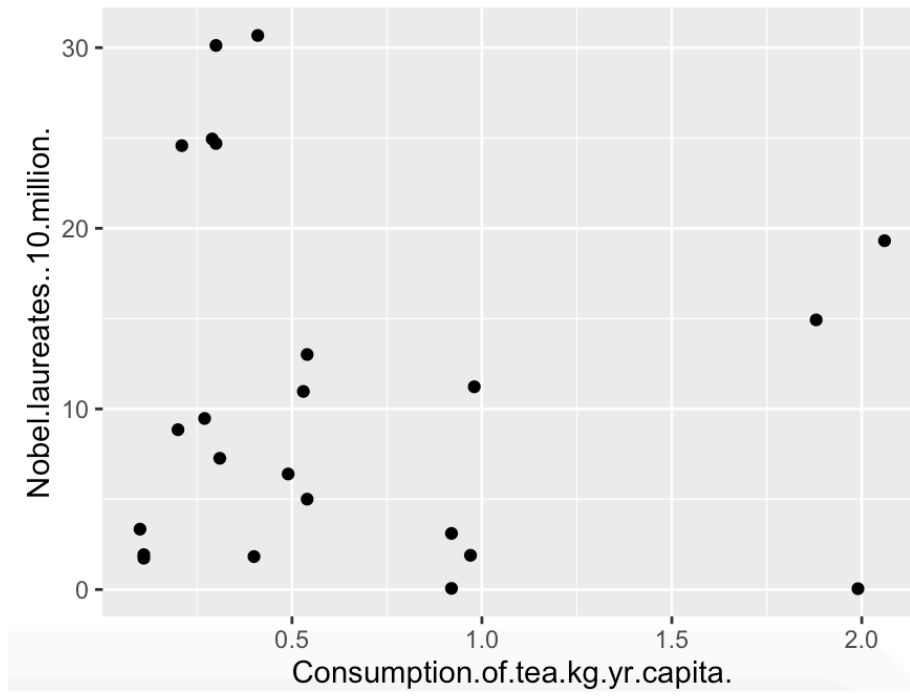
2.Chart

The following chart is the correlation between 'Chocolate consumption' and 'Nobel laureates'. The result is quite same as Franz's result.
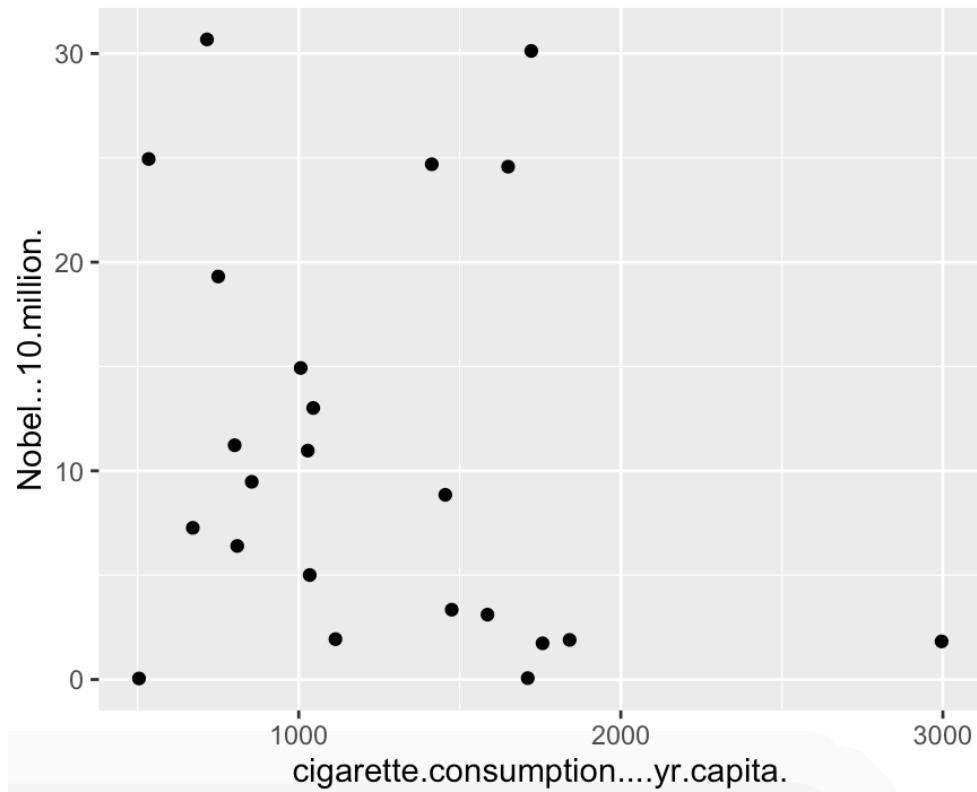

And then, we generate the charts which relate Nobel laureates to Tea consumption, Alcohol consumption and Cigarettes consumption.
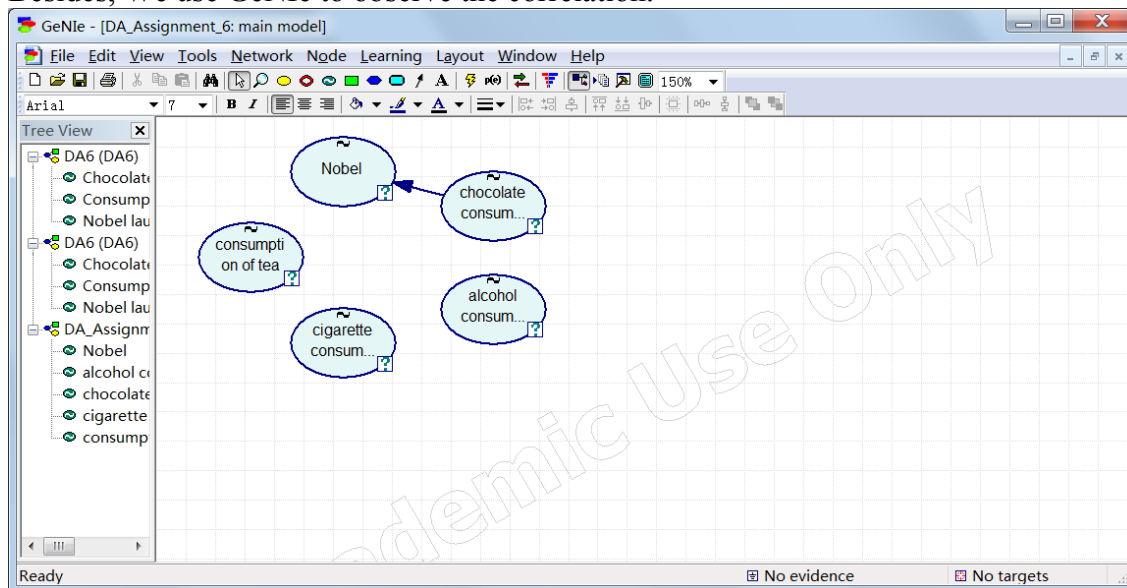
X: Tea consumption(kg/yr/capita), Alcohol consumption (litre/yr/capita), Cigarette consumption (/yr/capita)
Y: Nobel laureates(/10 million)

Besides, We use GeNIe to observe the correlation.

Nobel=2.66188*chocolate_consumption+Normal(-2.19229,6.42313)

According to the several charts above, we can make the following conclusions:

1. We can tell that chocolate consumption is correlated to the Nobel Laureates as is shown in the initial report.

2. The correlation between other factors and Nobel laureates is not very obvious. But when we remove several outliers, we can see that Nobel laureates and tea consumption as well as cigarattes are probably negatively related.

References
1. Chocolate Consumption: http://www.statista.com/statistics/262981/amount-of-chocolate-consumed-per-person-by-country/
2. Nobel: https://en.wikipedia.org/wiki/List_of_countries_by_Nobel_laureates_per_capita
3. Tea Consumption
   https://en.wikipedia.org/wiki/List_of_countries_by_tea_consumption_per_capita