

Data Analysis and Visualization

ACCB4002

Student: Iulian Raileanu

Lecturer: Anupam Mazumdar

1920709

Table of Contents:

Introduction.....	3
Dataset.....	4
Hypotheses.....	6
t-Test: Two-Sample Assuming Unequal Variances.....	7
t-Test $p=0.01$	7
t-Test $p=0.02$	8
t-Test $p=0.10$	9
Applications.....	10
Result.....	10
Conclusion.....	10
References.....	11

Introduction

This report shows the results of testing and comparing the information collected by BBC, all over United Kingdom. BBC said that the contamination of Corona Virus (COVID 19) of the infection number of White and BAME (Black, Asians and Minority Ethnic) peoples differ in United Kingdom and the changing and the speculation is made based on that information. This report contains the t-Tests for inconsistent changes which is to be performed on the information based. The result, applications and the charts and tables will be shown and discussed in this report.

Dataset

In a new BBC (United Kingdom) news, it has been distinguished that the contamination pace of Covid (Coronavirus) is changed between white individuals and BAME in the United Kingdom area. To recognize the proportion and the honesty of the discoveries, 50 diverse white and BAME patients have been chosen. All the patient's age is somewhere in the range of 20 and 55 years. In light of the information in the table underneath, you will see the result of the proportion of disease rate and honesty of the discoveries. The dataset used to make this report are shown in **Figure 1** and **Figure 2**

Patient no.	White	BAME	Patient no.	White	BAME
1	20	23	26	50	50
2	22	34	27	32	33
3	25	44	28	20	45
4	26	45	29	30	47
5	29	29	30	40	20
6	27	27	31	44	32
7	48	35	32	35	22
8	43	45	33	28	28
9	55	44	34	40	30
10	43	30	35	43	34
11	27	40	36	38	32
12	34	50	37	33	44
13	44	50	38	34	47
14	50	50	39	50	30
15	37	41	40	22	44
16	35	43	41	37	29
17	33	47	42	45	40
18	40	32	43	20	30
19	42	31	44	44	40
20	48	20	45	33	24
21	49	26	46	20	50
22	30	28	47	38	46
23	33	30	48	35	30
24	43	39	49	25	27
25	50	50	50	40	50

Figure 1.

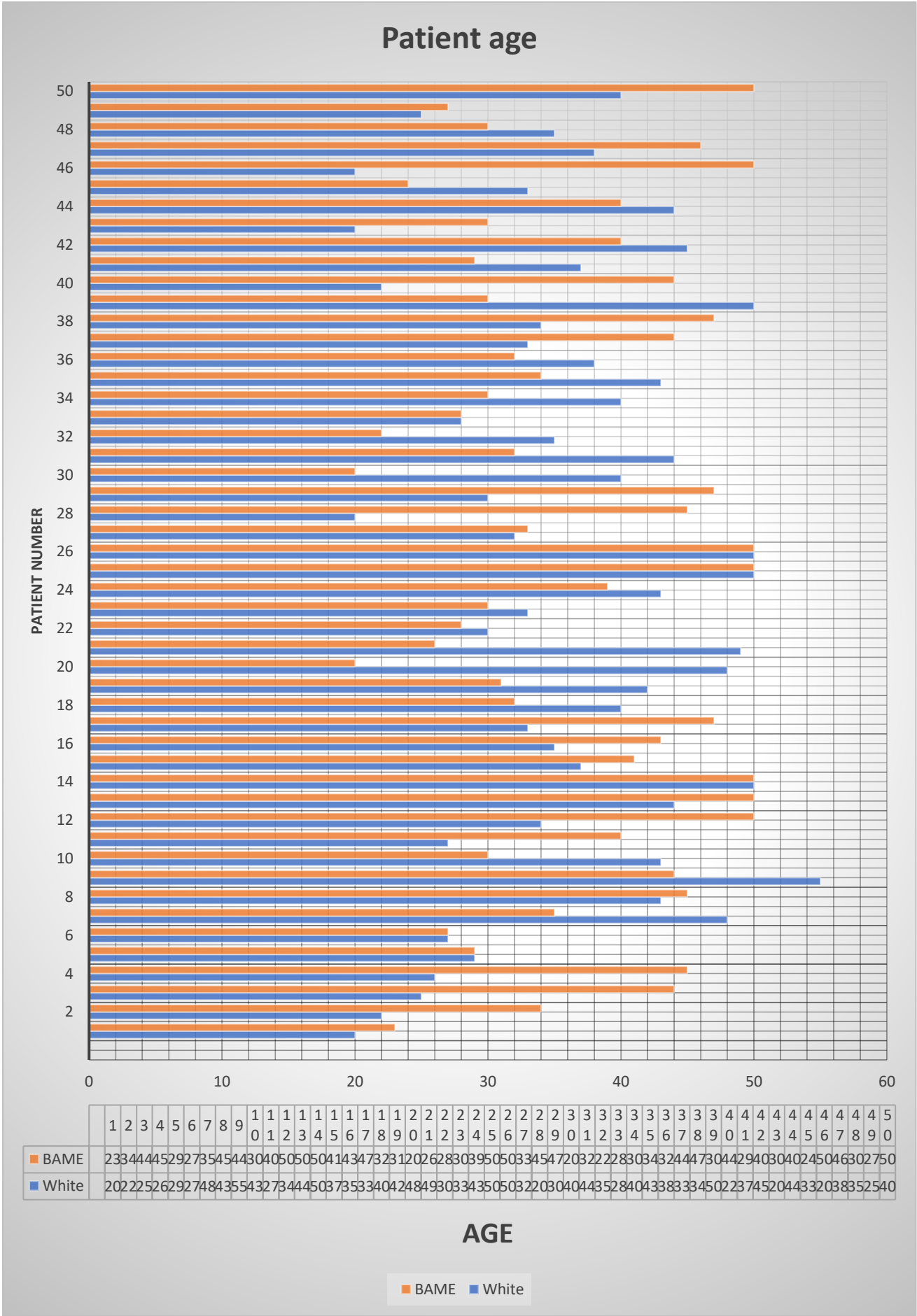


Figure 2

Hypotheses

A Hypothesis can be defined as an assumption or concession made for contention; an understanding of a functional circumstance or condition made as the ground for action; a provisional supposition made to draw and test its coherent or exact results; the predecessor provision of a restrictive assertion. (1)

In statistical hypothesis testing, we compare two of them, named **null hypothesis** and **alternative hypothesis**. According to Douglas G. Altman (2) in his book, the **null hypothesis** is the hypothesis that expresses that there is no connection between the phenomenon that is under investigation, or at least not of the result shown by the alternative hypothesis. The **alternative hypothesis** is the alternative to the null hypothesis. This hypothesis can take different forms, depending of the facts of the analyzed relation.

Null Hypothesis is noted with H_0 and it shows that there is no relation between the variables.

Alternative Hypothesis it is noted with H_1 and it shows the result that is expected.

t-Test represent a type of inferential statistic used to find out if there is a big difference between the means of two groups.(3) We have 3 types of t-Test:

- t-Test: Paired Two Sample for Mean
- t-Test: Two-Sample Assuming Equal Variance
- t-Test: Two-Sample Assuming Unequal Variance

t-Test:Two-Sample Assuming Unequal Variances:

with significance value $p=0.01$ *Figure2*

t-Test: Two-Sample Assuming Unequal Variances

	Variable 1	Variable 2
Mean	36.18	36.74
Variance	88.72204082	86.80857143
Observations	50	50
Hypothesized Mean Difference	0	
df	98	
t Stat	-0.298879822	
P(T<=t) one-tail	0.382832305	
t Critical one-tail	2.36500241	
P(T<=t) two-tail	0.76566461	
t Critical two-tail	2.626931096	

Figure 2

For the p-value of 0.01, the test is performed and discovered the result.

We choose from: **Null hypothesis H_0** that there is no significant difference between the mean of white patients and BAME patients.

Or **Alternative Hypothesis H_1** that there is a significant difference between the mean of white patients and BAME patients.

Null hypothesis H_0 can be accepted if:

- a. the value of t Stat is smaller than the individual value of t Critical or
- b. if the P-value is greater than the Alpha number which here we have it 0.01
- a. t stat is -0.29 (-0.30 rounded) and the t Critical one tail is 2.36 (2.37 rounded)
- b. P two value is 0.76 (0.77) and the significance value is 0.01
- a. Because of the results we can accept the **null hypothesis H_0**
- b. Because of the results we can accept the **null hypothesis H_0**

From both cases, we see that the **null hypothesis H_0** it is made as cannot be rejected

with significance value $p=0.02$ **Figure 3.**

t-Test: Two-Sample Assuming Unequal Variances 0.02		
	Variable 1	Variable 2
Mean	36.18	36.74
Variance	88.72204082	86.80857143
Observations	50	50
Hypothesized Mean Difference	0	
df	98	
t Stat	-0.298879822	
P(T<=t) one-tail	0.382832305	
t Critical one-tail	2.081444921	
P(T<=t) two-tail	0.76566461	
t Critical two-tail	2.36500241	

Figure 3

Null hypothesis H_0 that there is no significant difference between the mean of white patients and BAME patients.

Alternative Hypothesis H_1 that there is a significant difference between the mean of white patients and BAME patients.

- T Stat value -0.30 rounded and T Critical one tail 2.08
- P two value is 0.77 rounded and significance value is 0.02

Here both answers are showing us again that the **null hypothesis can not be rejected**

With significance value $p=0.10$ *Figure 4.*

t-Test: Two-Sample Assuming Unequal Variances 0.1		
	Variable 1	Variable 2
Mean	36.18	36.74
Variance	88.72204082	86.80857143
Observations	50	50
Hypothesized Mean Difference	0	
Df	98	
t Stat	-0.298879822	
P(T<=t) one-tail	0.382832305	
t Critical one-tail	1.290249904	
P(T<=t) two-tail	0.76566461	
t Critical two-tail	1.660551217	

Figure 4

Null hypothesis H_0 that there is no significant difference between the mean of white patients and BAME patients.

Alternative Hypothesis H_1 that there is a significant difference between the mean of white patients and BAME patients.

- tStat -0.30 rounded and t Critical is 1.29
- P value is 0.76 and significance value is 0.10

After these result we can not reject the **null hypothesis H_0**

Applications

T-Tests can be utilized to compare the mean of two samples, dependent or independent, can be utilized to decide if the example means is not the same as the expected to be mean and also can be used to decide the certainty interval for an example mean ([4](#))

Result

From the results of the t-Tests with the significance numbers of 0.01, 0.02 and 0.10 we can see that the Null hypotheses can not be dismissed in all cases so the supposition which was made as the contamination with Corona Virus rates in the BAME and White patients is fluctuating is valid.

Conclusion

In this report, the t-Tests it was made utilizing the Microsoft Excel formulas, the Data of patients finded by BBC, of persons from United Kingdom, by splitting in two parts, white and BAME, the people infected with Covid 19.

References

1. Merriam-webster <https://www.merriam-webster.com/dictionary/hypothesis>
2. Douglas G. Altman PRACTICAL STATISTICS FOR MEDICAL RESEARCH 1991
3. <https://www.investopedia.com/terms/t/t-test.asp>
4. <https://microbenotes.com/t-test/#t-test-applications>