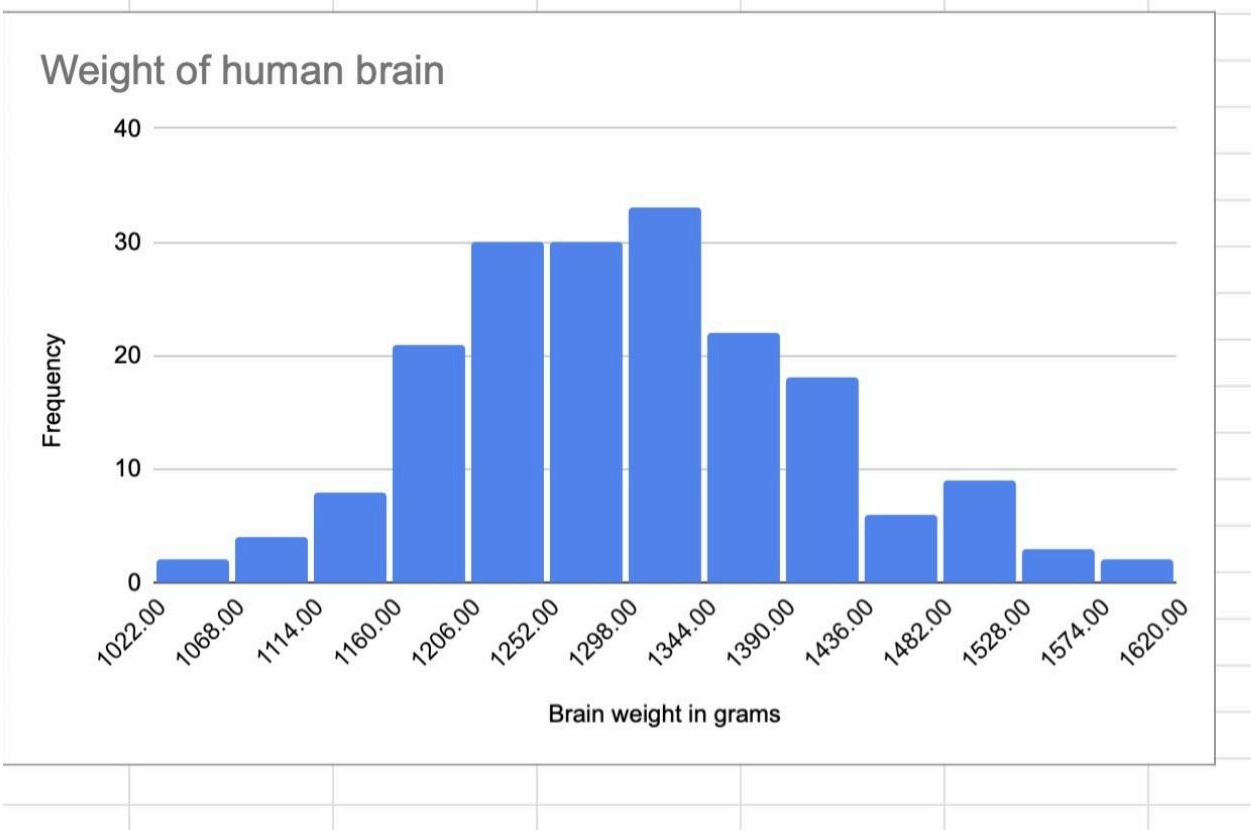
1. Firstly, I started to search for the data on the credible website Kaggle.com, which collected all datasets from special researchers, other academic institutions, and governmental reports. I wanted to find the data that would be normally distributed. Consequently, I looked up into the data's histogram with a bell-shaped figure, which meant that normally distributed data has a symmetrical skewness. Moreover, I found out that height and weight of people tend to be normally distributed. Therefore, I found the data from the primary research by Hartmann et al.(1994) named "Normal weight of the brain in adults in relation to age, sex, body height and weight". This research conducted a study by using data from people without brian disease in Germany to find a normal brain weight. The data is about the human brain's weight in grams. I chose 189 values from 237, since it clearly showed in histogram a normal distribution by being symmetric. Moreover, the dataset was divided into 4 sections encompassing age range, head size, brain weight and gender. Here is a link for the datasets: https://www.kaggle.com/datasets/anubhabswain/brain-weight-in-humans
   Moreover, there is a link for the research : https://pubmed.ncbi.nlm.nih.gov/8072950/
2. I hypothesize that the dataset of human brian weight is normally distributed and can be accepted since the test statistic (9.072) can be less than 1%, 2.5% and 5% levels of significance .
3. I found out that Pearson's coefficient of skewness was 0.145, which meant that data was closer to 0, making the data more symmetrical and normally distributed. It is also positively skewed. This means that most of the weights of the brain were relatively low, but there was a tail of big weights of the brain. It means that the data might be normally distributed since the graphic of this data is more likely to be bell-shaped.

mean	1302.407
median	1297.000
st.dev	112.178
skewness	0.145

4. The histogram bucket size was 46, since minimum of the data was 1022 and maximum of the data was 1620. The difference was 598, which made me choose the class width 46, and the number of the bars was 13. Moving forward, this histogram is slightly symmetrical and resembles a bell, even though it has a very small right tail. This can indicate that the data can be considered as normally distributed , since the histogram is mostly symmetrical.



For normally distributed data it is known that 68.3% of the data lies on the interval between (mean-standard deviation, mean+standard deviation) and 95.5% of the data lies on (mean-2*standard deviation, mean+2*standard deviation) and 99.7% of the data lies on between (mean-3*standard deviation, mean+3*standard deviation). However, in my case the data did not perfectly match these criterias. This highlights that the data is not ideally normally distributed, even though it is similar for criterias and there are slight differences. Also, it can be concluded that the data can have a potential to be normally distributed. In addition to that, these percentages show that the hypothesis about the dataset being normally distributed can be proven. Because these percentages are very close to the criterias, which indicates that even if the dataset is not ideally normally distributed, it has a potential to be normally distributed.

left	right	frequency	percentage	
1190.230	1414.585	129	68.25396825	(1*st dev)
1078.052	1526.763	178	94.17989418	(2*st dev)
965.875	1638.940	189	100	(3*st dev)

5. The chi squared table:

Intervals		Observed frequency	standardized intervals		P(Z<lcb)	P(Z<ucb)	Expected frequency
lcb	ucb		lcb	ucb			
1022	1068	2	0.000	-2.090	0.000	0.018	3.464
1068	1114	4	-2.090	-1.680	0.018	0.047	5.329
1114	1160	8	-1.680	-1.269	0.047	0.102	10.511
1160	1206	21	-1.269	-0.859	0.102	0.195	17.562
1206	1252	30	-0.859	-0.449	0.195	0.327	24.860
1252	1298	30	-0.449	-0.039	0.327	0.484	29.813
1298	1344	33	-0.039	0.371	0.484	0.645	30.291
1344	1390	22	0.371	0.781	0.645	0.783	26.073
1390	1436	18	0.781	1.191	0.783	0.883	19.014
1436	1482	6	1.191	1.601	0.883	0.945	11.747
1482	1528	9	1.601	2.011	0.945	0.978	6.148
1528	1574	3	2.011	2.421	0.978	0.992	2.726
1574	1620	3	2.421	1.000	0.992	1.000	1.462
		189					189.000

Expected frequency	Observed frequency	(O-E)^2/E
8.793	6	0.887
10.511	8	0.600
17.562	21	0.673
24.860	30	1.063
29.813	30	0.001
30.291	33	0.242
26.073	22	0.636
19.014	18	0.054
11.747	6	2.812
10.337	15	2.104
189.000	189	9.072

6. Conclusion: The degree of freedom was 7, which was calculated by removing a number of estimated parameters(3) from the number of cells after combining (10). Consequently, the critical value for 1% level of significance was 18.48, and for the 2.5% level of significance was 16.01, and also critical value for 5% level of significance was 14.07. All of these critical values were more than 9.072, which indicated that the hypothesis can be accepted and the data can be considered as normally distributed.

Sig level	Critical value	
1%	18.48	Accept the hypothesis since O21>L32
2.50%	16.01	Accept the hyotheisis since O22>L32
5%	14.07	Accept the hypothesis since O23>L32