

Resultados - Box-plot

```
In [1]: 1 install.packages("readxl")
        2 install.packages("ggpubr")
        3 install.packages("xlsx")

Installing package into '/home/fsa/R/x86_64-pc-linux-gnu-library/3.6'
(as 'lib' is unspecified)
Installing package into '/home/fsa/R/x86_64-pc-linux-gnu-library/3.6'
(as 'lib' is unspecified)
Installing package into '/home/fsa/R/x86_64-pc-linux-gnu-library/3.6'
(as 'lib' is unspecified)
```

```
In [2]: 1 library("readxl")
        2 library(writexl)
        3 library(xlsx)
        4 library(ggthemes)
        5 library(ggplot2)
        6 library(plotly)
        7 library(dplyr)
        8 library(ggpubr)
```

Attaching package: 'plotly'

The following object is masked from 'package:ggplot2':

last_plot

The following object is masked from 'package:stats':

filter

The following object is masked from 'package:graphics':

layout

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

Loading required package: magrittr

```
In [3]: 1 dados <- read_excel("Economics.xlsx", sheet = "n-Value_EDITED")
```

New names:

```
* `` -> ...3
* `` -> ...4
* `` -> ...5
* `` -> ...6
* `` -> ...7
* ... and 19 more problems
```

In [4]:

	1	dados								
	Grupos	BD	...3	...4	...5	...6	...7	...8	HC	...
	Alpha-2-antiplasmin	16.78596	18.05648	15.69936	14.44074	16.72014	16.60659	16.97074	18.47524	16.422
	Serum amyloid P-component	16.24382	17.14231	15.76176	17.30153	17.00755	17.28555	16.35430	16.70063	15.830
	Coagulation factor XIII B chain	16.02046	15.79020	17.31857	15.55219	16.12434	15.96810	15.68294	15.90616	16.619
	Inter-alpha-trypsin inhibitor heavy chain H4	20.96332	20.67702	20.54189	20.74706	21.04347	21.29603	20.57089	21.89078	21.349
	Apolipoprotein A-II	22.98031	21.90709	24.07006	23.03159	23.10642	22.70571	23.16249	22.83040	22.576
	Platelet basic protein	17.94575	18.26109	16.21544	19.31816	18.04383	18.61956	17.18911	18.18979	18.690
	Carboxypeptidase N subunit 2	16.21534	16.30048	15.61976	16.16632	17.03585	16.39178	15.98471	16.69411	16.230

Criação dos Data-Frames dos grupos de doenças mentais

In [5]:

```
1 bipolar_data <- data.frame("x1" = dados$BD,
2                             "x2" = dados$...3,
3                             "x3" = dados$...4,
4                             "x4" = dados$...5,
5                             "x5" = dados$...6,
6                             "x6" = dados$...7,
7                             "x7" = dados$...8)
8
9 row.names(bipolar_data) <- dados$Grupos
```

In [6]:

1	bipolar_data
2	

	x1	x2	x3	x4	x5	x6	x7
Alpha-2-antiplasmin	16.78596	18.05648	15.69936	14.44074	16.72014	16.60659	16.97074
Serum amyloid P-component	16.24382	17.14231	15.76176	17.30153	17.00755	17.28555	16.35430
Coagulation factor XIII B chain	16.02046	15.79020	17.31857	15.55219	16.12434	15.96810	15.68294
Inter-alpha-trypsin inhibitor heavy chain H4	20.96332	20.67702	20.54189	20.74706	21.04347	21.29603	20.57089
Apolipoprotein A-II	22.98031	21.90709	24.07006	23.03159	23.10642	22.70571	23.16249
Platelet basic protein	17.94575	18.26109	16.21544	19.31816	18.04383	18.61956	17.18911
Carboxypeptidase N subunit 2	16.21534	16.30048	15.61976	16.16632	17.03585	16.39178	15.98471
Haptoglobin	23.58532	23.53362	23.59651	23.84221	23.31645	22.49644	23.87549
Glutathione peroxidase 3	15.77471	14.68699	15.43785	14.89562	15.32892	15.54104	15.24659
Heparin cofactor 2	18.05944	18.72910	17.40116	17.94318	18.60675	17.94392	18.93163
Alpha-1-antitrypsin	23.89097	23.85785	23.68260	24.42137	23.84479	24.52486	23.16967
Apolipoprotein A-I	24.96596	24.09901	24.81924	24.01657	24.51653	24.61116	24.28320
Vitamin K-dependent protein S	17.76578	18.45004	18.42521	18.06287	18.06055	18.37954	18.24776
Alpha-2-HS-glycoprotein	21.24383	21.89538	21.94575	22.34889	21.82881	22.31915	21.06151
Fibronectin 1	21.05762	21.19298	19.83189	20.48837	20.95858	21.10412	20.57228
Apolipoprotein D	16.70523	16.24299	16.76552	17.33429	16.88178	17.82196	19.60695
Apolipoprotein A-IV	20.59732	20.80806	21.12723	20.82734	20.50475	21.24806	20.79463
Prothrombin	20.18930	21.10310	22.07988	20.16103	20.12578	20.87620	20.29480
Alpha-1-acid glycoprotein 1	21.75764	22.40600	21.92278	21.73913	22.52306	22.28915	22.39223
Complement factor H	21.97197	22.17329	22.40901	22.33451	22.56440	22.62723	22.15755
Haptoglobin-related protein	18.71482	18.62325	20.45503	18.82258	18.92164	19.92687	18.66518
Apolipoprotein B-100	23.40510	23.71712	23.93256	23.21614	23.63123	23.49777	23.23631
Plasma protease C1 inhibitor	19.23582	19.30247	21.40561	19.54958	19.08672	19.52396	19.82627
Pregnancy zone protein	15.32372	14.51039	14.95665	13.45529	16.59731	16.37533	12.60890
Hemopexin	23.59423	23.53600	23.93643	23.52480	23.61852	23.83123	22.98944
Complement C8 alpha chain	16.20149	16.26224	16.77483	16.44121	16.71585	16.54480	17.64343
Complement component C8 gamma chain	17.62649	18.19816	17.79127	18.47183	18.51768	18.22515	17.46818
IgJ chain	16.96251	15.64740	16.36293	16.06842	16.47856	16.61884	17.40333
Complement factor I	17.90360	17.99937	18.41055	18.24665	18.56852	18.96065	18.01900
Clusterin	20.60159	20.85780	21.23316	21.76255	21.02513	21.01134	20.42726
Butyrylcholinesterase	16.01110	16.82799	16.42028	16.33703	16.45189	16.58939	16.15284
Coagulation factor XII	18.37420	18.99604	15.38323	18.35990	18.52187	17.78788	17.03843
C4b-binding protein alpha chain	20.25753	20.53602	18.30768	20.76061	20.50543	20.81613	20.11337
Apolipoprotein M	16.04367	16.79538	16.03704	14.91625	15.95374	15.72353	16.45330

```
In [7]: 1 controle_data <- data.frame("x1" = dados$HC,
2                                     "x2" = dados$...10,
3                                     "x3" = dados$...11,
4                                     "x4" = dados$...12,
5                                     "x5" = dados$...13,
6                                     "x6" = dados$...14,
7                                     "x7" = dados$...15,
8                                     "x8" = dados$...16,
9                                     "x9" = dados$...17,
10                                    "x10" = dados$...18,
11                                    "x11" = dados$...19,
12                                    "x12" = dados$...20,
13                                    "x13" = dados$...21)
14
15 row_names(controle_data) <- dados$Grupos
```

```
In [8]: 1 controle_data
```

	x1	x2	x3	x4	x5	x6	x7	x8	
Alpha-2-antiplasmin	18.47524	16.42244	16.16857	17.58701	20.27185	18.03676	18.40736	19.17959	17.21
Serum amyloid P-component	16.70063	15.83029	18.22680	17.75193	18.32364	16.63962	16.56577	17.25248	17.11
Coagulation factor XIII B chain	15.90616	16.61999	16.52595	17.14049	15.90221	15.48963	16.82588	16.24390	15.71
Inter-alpha-trypsin inhibitor heavy chain H4	21.89078	21.34965	21.53867	21.20729	20.77427	22.25664	20.42859	20.60540	21.61
Apolipoprotein A-II	22.83040	22.57630	21.65949	19.82833	19.97113	22.89754	23.34001	22.03043	22.21
Platelet basic protein	18.18979	18.69074	18.84650	16.09975	17.44581	18.84739	14.16608	18.36863	17.81
Carboxypeptidase N subunit 2	16.69411	16.23013	15.44928	16.50285	15.37819	17.14490	19.25403	16.11405	17.01

```
In [9]: 1 esquizo_data <- data.frame("x1" = dados$SCZ,
2                                     "x2" = dados$...23,
3                                     "x3" = dados$...24,
4                                     "x4" = dados$...25,
5                                     "x5" = dados$...26,
6                                     "x6" = dados$...27,
7                                     "x7" = dados$...28)
8
9 row_names(esquizo_data) <- dados$Grupos
```

In [10]:

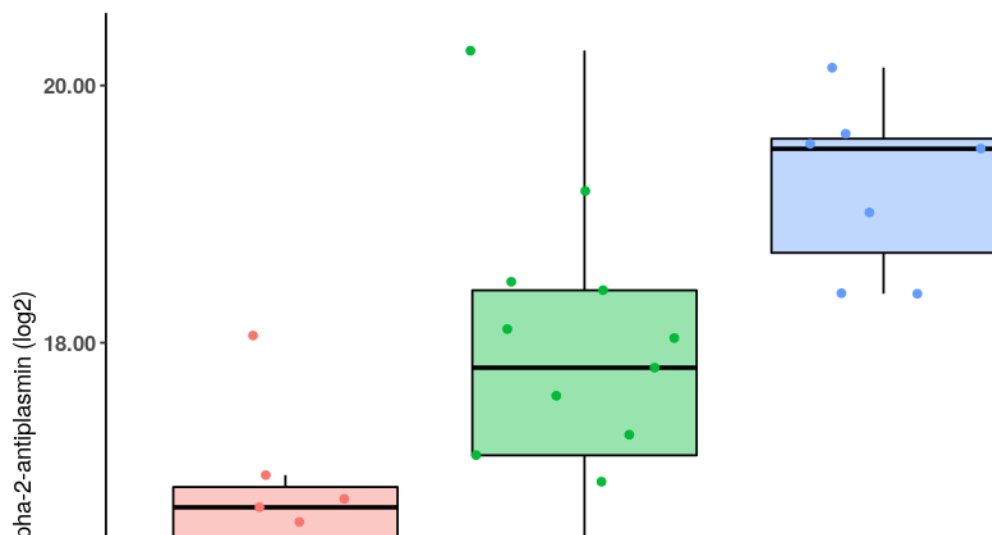
1 esquizo_data

	x1	x2	x3	x4	x5	x6	x7
Alpha-2-antiplasmin	19.54764	18.38545	20.13884	19.62501	19.50798	18.38119	19.01258
Serum amyloid P-component	18.52573	17.77069	18.16074	18.12999	18.44107	17.22703	18.75418
Coagulation factor XIII B chain	16.93041	16.85404	17.36170	17.03628	16.76046	16.69860	16.68099
Inter-alpha-trypsin inhibitor heavy chain H4	22.83148	21.84487	21.31529	21.30084	21.68496	21.60612	21.75666
Apolipoprotein A-II	20.93524	19.40811	21.64042	21.53645	21.87335	20.96346	22.27807
Platelet basic protein	19.62776	20.61477	19.18171	19.13976	19.06745	19.46702	18.18490
Carboxypeptidase N subunit 2	17.90006	16.64118	17.71773	17.48030	17.48093	17.06855	16.80663
Haptoglobin	25.82350	25.93373	23.73624	23.66590	24.22920	23.70588	23.50628
Glutathione peroxidase 3	15.03381	15.58317	15.30930	14.44657	13.99418	15.28435	14.44443
Heparin cofactor 2	19.94692	19.32518	19.11589	19.17616	19.15889	18.48625	19.39475
Alpha-1-antitrypsin	25.17798	25.41806	24.16928	23.97710	24.55104	24.27327	24.03078
Apolipoprotein A-I	23.76871	24.34449	24.08221	24.13156	23.34597	24.21334	23.49558
Vitamin K-dependent protein S	19.50219	18.09151	18.61310	18.84500	19.01838	18.54359	19.36072
Alpha-2-HS-glycoprotein	22.37579	22.97463	21.52032	21.81026	21.46309	21.65958	21.79812
Fibronectin 1	21.04046	21.86007	21.26023	21.47556	20.82951	21.61279	21.95397
Apolipoprotein D	14.86873	15.83061	16.23244	16.10617	13.92333	16.10607	14.54472
Apolipoprotein A-IV	21.05470	19.64688	21.10463	20.54217	20.50611	21.02763	18.90448
Prothrombin	21.66371	20.97456	21.29855	21.32036	21.40145	21.31761	21.23199
Alpha-1-acid glycoprotein 1	23.24644	22.15814	22.11006	22.40671	23.16759	21.79812	22.96739
Complement factor H	22.49380	22.51396	22.37162	22.48134	22.51060	22.70841	22.64889
Haptoglobin-related protein	20.21050	18.67097	20.26671	20.35434	20.16582	20.26408	18.96110
Apolipoprotein B-100	24.03053	23.59012	23.99697	24.15290	23.34055	24.24213	23.74212
Plasma protease C1 inhibitor	19.45210	19.06329	18.96531	18.52929	18.68815	19.17553	18.43090
Pregnancy zone protein	16.55499	21.98384	13.80312	17.69534	13.65765	13.91933	15.99612
Hemopexin	23.55684	22.67532	23.13297	23.00137	23.26227	23.46849	22.87245
Complement C8 alpha chain	17.43949	16.27132	17.62092	17.70286	17.85075	17.26114	18.12182
Complement component C8 gamma chain	18.44330	18.19215	18.92141	19.03408	18.71112	19.00330	18.32246
IgJ chain	14.64135	17.13711	15.56084	15.61876	16.04097	16.29873	15.57406
Complement factor I	18.76325	19.03869	18.40715	18.57093	19.17750	19.03158	19.60462
Clusterin	21.45036	21.79891	21.30145	21.42157	21.39305	21.64619	21.43206
Butyrylcholinesterase	15.60322	16.64202	16.56254	18.69636	17.71827	16.91684	17.70725
Coagulation factor XII	19.23093	17.86653	18.13400	19.17809	18.42366	18.63803	19.16391
C4b-binding protein alpha chain	21.63139	20.82314	20.14644	21.05676	21.20538	20.78796	21.38943
Apolipoprotein M	15.62827	12.58949	16.32119	17.12212	16.73886	16.22865	18.34130

```

In [11]: 1 for(i in c(1:length(dados$Grupos))){
2
3     controle <- controle_data %>% slice(i) %>% t() %>% as.data.frame() %>%
4     esquizo <- esquizo_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
5     bipolar <- bipolar_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
6     group <- rbind(controle, esquizo,bipolar) %>% rename(value = V1)
7
8     plot <- ggplot(data = group,
9                   aes(x = label,
10                      y = value,
11                      #alpha = 0.2,
12                      colour = as.factor(label))) +
13
14     geom_boxplot(aes(fill = as.factor(label)),
15                 outlier.colour = "white",
16                 alpha = 0.4,
17                 colour = "black") +
18
19     geom_jitter() +
20
21     theme(legend.position = "none")+
22
23     labs(x= "Condition",
24          y= paste(dados$Grupos[i],
25                  "(log2)"),
26          title= NULL) +
27
28     theme(
29         axis.line = element_line(colour = "black"),
30         panel.grid.major = element_blank(),
31         panel.grid.minor = element_blank(),
32         panel.border = element_blank(),
33         panel.background = element_blank(),
34         axis.text.x = element_text(angle = 90, size = 13, face = "bold"),
35         axis.text.y = element_text(size = 10,face = "bold")) +
36
37     scale_y_continuous(
38         labels = scales::number_format(accuracy = 0.01,
39         decimal.mark = '.'))
40
41     print(plot)
42 }
43
44

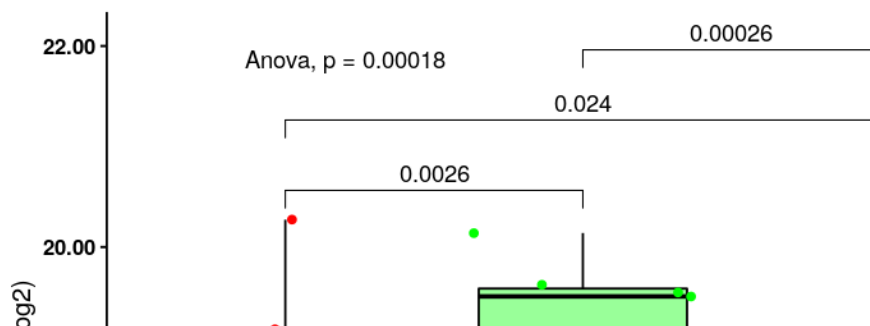
```



```

In [18]: 1 for(i in c(1:length(dados$Grupos))){
2
3   controle <- controle_data %>% slice(i) %>% t() %>% as.data.frame() %>%
4   esquizo <- esquizo_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
5   bipolar <- bipolar_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
6   group <- rbind(controle, esquizo,bipolar) %>% rename(value = V1)
7
8
9   my_comparisons <- list( c("Control", "Schizophrenia"),
10                          c("Control", "Bipolar"),
11                          c("Bipolar", "Schizophrenia"))
12
13
14   p <- ggboxplot(group,
15                 x = "label",
16                 y = "value",
17                 color = "black",
18                 fill = "label",
19                 alpha = 0.4,
20                 outlier.colour = "white",
21                 palette = c("red", "green", "blue")) +
22
23
24
25   stat_compare_means(comparisons = my_comparisons,
26                     method = "t.test") +
27
28   stat_compare_means(method = "anova",
29                     label.y = max(group$value) + 1.6) +
30
31   geom_jitter(aes(color = label)) +
32
33
34   theme(legend.position = "none") +
35
36   labs(x= "Condition",
37        y= paste(dados$Grupos[i], "(log2)"),
38        title= NULL) +
39
40   theme(
41     axis.line = element_line(colour = "black"),
42     panel.grid.major = element_blank(),
43     panel.grid.minor = element_blank(),
44     panel.border = element_blank(),
45     panel.background = element_blank(),
46     axis.text.x = element_text(angle = 90, size = 12, face = "bold"),
47     axis.text.y = element_text(size = 10, face = "bold")) +
48
49   scale_y_continuous(labels = scales::number_format(accuracy = 0.01,
50                                                       decimal.mark = '.'))
51
52   print(p)
53
54 }
55

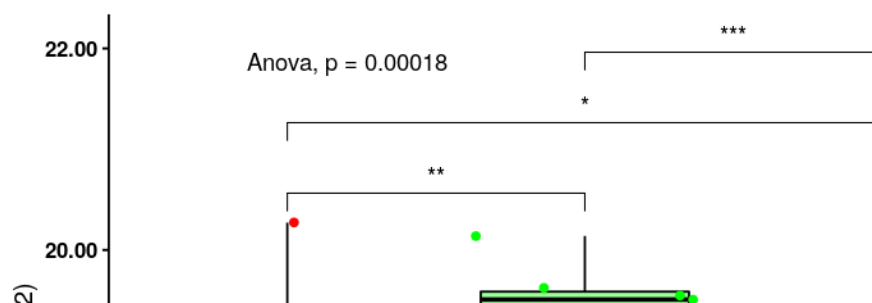
```



```

In [20]: 1 for(i in c(1:length(dados$Grupos))){
2
3   controle <- controle_data %>% slice(i) %>% t() %>% as.data.frame() %>%
4   esquizo <- esquizo_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
5   bipolar <- bipolar_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
6   group <- rbind(controle, esquizo,bipolar) %>% rename(value = V1)
7
8
9   my_comparisons <- list( c("Control", "Schizophrenia"),
10                          c("Control", "Bipolar"),
11                          c("Bipolar", "Schizophrenia"))
12
13
14   p <- ggboxplot(group,
15                 x = "label",
16                 y = "value",
17                 color = "black",
18                 fill = "label",
19                 alpha = 0.4,
20                 outlier.colour = "white",
21                 palette = c("red", "green", "blue")) +
22
23
24
25   stat_compare_means(comparisons = my_comparisons,
26                     label = "p.signif", method = "t.test",
27                     ref.group = "0.5")+
28
29   stat_compare_means(method = "anova",
30                     label.y = max(group$value) + 1.6) +
31
32   geom_jitter(aes(color = label)) +
33
34
35   theme(legend.position = "none") +
36
37   labs(x= "Condition",
38        y= paste(dados$Grupos[i], "(log2)"),
39        title= NULL) +
40
41   theme(
42     axis.line = element_line(colour = "black"),
43     panel.grid.major = element_blank(),
44     panel.grid.minor = element_blank(),
45     panel.border = element_blank(),
46     panel.background = element_blank(),
47     axis.text.x = element_text(angle = 90, size = 12, face = "bold"),
48     axis.text.y = element_text(size = 10, face = "bold")) +
49
50   scale_y_continuous(labels = scales::number_format(accuracy = 0.01,
51                                                       decimal.mark = '.'))
52
53   print(p)
54
55 }
56

```




```
In [ ]: 
```

```

In [13]: 1 for(i in c(1:length(dados$Grupos))){
2
3     controle <- controle_data %>% slice(i) %>% t() %>% as.data.frame() %>%
4     esquizo <- esquizo_data %>% slice(i) %>% t() %>% as.data.frame() %>%
5     bipolar <- bipolar_data %>% slice(i) %>% t() %>% as.data.frame() %>%
6
7     controle_esquizo <- rbind(controle,esquizo) %>% mutate(group = "Control x Schizophrenia")
8     controle_bipolar <- rbind(controle,bipolar) %>% mutate(group = "Control x Bipolar")
9     esquizo_bipolar <- rbind(esquizo,bipolar) %>% mutate(group = "Schizophrenia x Bipolar")
10
11
12
13
14     group <- rbind(controle_esquizo, controle_bipolar,esquizo_bipolar) %>%
15
16
17     my_comparisons <- list( c("Control", "Schizophrenia"),
18                             c("Control","Bipolar"),
19                             c("Bipolar","Schizophrenia"))
20
21     labels <- c(Control_Esquizo = "Control x Schizophrenia ",
22                 Control_Bipolar = "Control x Bipolar",
23                 Esquizo_Bipolar = "Schizophrenia x Bipolar")
24
25     p <- ggboxplot(group,
26                    x = "label",
27                    y = "value",
28                    color = "black",
29                    fill = "label",
30                    alpha = 0.4,
31                    palette = c("red","green","blue"),
32                    short.panel.labs = TRUE,
33                    outlier.colour = "white") +
34
35     facet_grid(group~.,drop=TRUE, labeller=labeller(group = labels)) +
36
37     stat_compare_means(method = "t.test",label.y = max(group$value) + 1.6)
38
39
40     geom_jitter(aes(color = label)) +
41
42
43     theme(legend.position = "none") +
44
45     labs(x= "Condition",
46          y= paste(dados$Grupos[i], "(log2)"),
47          title= NULL) +
48
49     theme(
50         axis.line = element_line(colour = "black"),
51         panel.grid.major = element_blank(),
52         panel.grid.minor = element_blank(),
53         panel.border = element_blank(),
54         panel.background = element_blank(),
55         axis.text.x = element_text(angle = 90, size = 12, face = "bold"),
56         axis.text.y = element_text(size = 10,face = "bold")) +
57
58     scale_y_continuous(labels = scales::number_format(accuracy = 0.01,
59                                                         decimal.mark = '.'))
60
61     print(p)
62
63 }
64

```

22.00

T-test, p = 0.024

22.00

•



In [14]:

```

1 controle_esquizo <- double(26)
2 controle_bipolar <- double(26)
3 bipolar_esquizo <- double(26)
4 proteina <- character(1)
5 anova <- double(26)
6
7
8 for(i in c(1:length(dados$Grupos))) {
9
10     controle <- controle_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
11     esquizo <- esquizo_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
12     bipolar <- bipolar_data %>% slice(i) %>% t() %>% as.data.frame() %>% n
13
14     proteina[i] <- dados$Grupos[i]
15     controle_esquizo[i] <- t.test(controle$V1,esquizo$V1)$p.value
16     controle_bipolar[i] <- t.test(controle$V1,bipolar$V1)$p.value
17     bipolar_esquizo[i] <- t.test(bipolar$V1,esquizo$V1)$p.value
18
19     vetor_anova <- c(controle$V1,esquizo$V1,bipolar$V1)
20
21     grupo_anova <- c(rep("control",length(controle$V1)),
22                     rep("esquizo",length(esquizo$V1)),
23                     rep("bipolar",length(bipolar$V1)))
24
25     anova[i] <- oneway.test(vetor_anova~grupo_anova,var.equal = TRUE)$p.val
26 }
27
28 tabela_consolidada <- data.frame("Proteina" = proteina,
29                                 "Controle_Esquizo" = controle_esquizo,
30                                 "Controle_Bipolar" = controle_bipolar,
31                                 "Bipolar_Esquizo" = bipolar_esquizo,
32                                 "Anova" = anova,
33                                 stringsAsFactors = FALSE)
34
35 tabela_consolidada <- tabela_consolidada %>% mutate("Anova p < 0.05" =
36             if_else(tabela_consolidada$Anova < 0.05,
37                     "True",
38                     "False"))
39
40 tabela_consolidada
41
42 write_xlsx(tabela_consolidada,
43            "tabela_consolidada.xlsx")

```

Proteina	Controle_Esquizo	Controle_Bipolar	Bipolar_Esquizo	Anova	Anova p < 0.05
Alpha-2-antiplasmin	0.0026028138	0.02380178	0.0002644458	0.0001809731	True
Serum amyloid P-component	0.0011456096	0.33705139	0.0005349176	0.0009908256	True
Coagulation factor XIII B chain	0.0002336652	0.82901762	0.0082396824	0.0038005296	True
Inter-alpha-trypsin inhibitor heavy chain H4	0.0843512321	0.01516531	0.0021977972	0.0042442038	True
Apolipoprotein A-II	0.0597425491	0.07455458	0.0018661150	0.0124932205	True
Platelet basic protein	0.0041697405	0.81173060	0.0130913399	0.0231527029	True
Carboxypeptidase N subunit 2	0.0280762696	0.37646901	0.0009305440	0.0379986450	True
Haptoglobin	0.1187300067	0.58498342	0.0695506540	0.0634225861	False
Glutathione peroxidase 3	0.0073562020	0.04910345	0.1539964997	0.0173935188	True
Heparin cofactor 2	0.3357906385	0.02942190	0.0027116405	0.0259685664	True
Alpha-2-macroglobulin	0.0050000000	0.00000000	0.00000000	0.00000000	True

