



POLITEKNIK
ELEKTRONIKA
NEGERI SURABAYA

LAPORAN TUGAS AKHIR

**Praktikum Statistika Dasar |
Statistik Non Parametrik**

DESEMBER 2021

DIBIMBING OLEH

Pak Ronny Susetyoko S.Si., M.Si.
Dosen Pengampu Mata Kuliah
Statistika
(197112111995011001)

DISUSUN OLEH

Muhammad Dzalhaqi
D-4 Sains Data terapan
(3321600023)

Statistika Non-Parametrik Menggunakan R Shiny

1 Pendahuluan

1.1 Latar Belakang

Data scientist adalah seorang yang memiliki kemampuan statistika lebih baik dibandingkan dengan para *Software Engineer* lain dan memiliki keahlian Programming yang lebih baik dibandingkan dengan para statiskawan lain. Oleh karena itu seorang yang mempelajari ilmu data haruslah memiliki kompetensi yang mumpuni dari segi ilmu komputer dan statistika. Salah satu topik statistika yang wajib dikuasai oleh seorang *data scientist* adalah statistik non parametrik. Pada topik tersebut terdapat beberapa pengujian dengan metode bebas-distribusi yang mana hal tersebut menghilangkan asumsi bahwa data yang diuji berdistribusi normal kecuali barangkali distribusi data tersebut merupakan distribusi kontinyu. Hal inilah yang kemudian mendasari pembuatan modul interaktif sederhana terkait pengujian statistik non-parametrik berbasis web menggunakan R Shiny.

1.2 Rumusan Masalah

1. Bagaimana cara mengetahui, memahami, dan menerapkan proses pengujian data menggunakan metode statistik non-parametrik melalui modul interaktif berbasis web menggunakan R Shiny?
2. Bagaimana menerapkan visualisasi data yang tepat berdasarkan tipe data dan metode pengujian statistik non parametrik?

1.3 Tujuan

1. memberikan kemudahan dalam mengetahui, mempelajari serta mengaplikasikan proses pengujian data menggunakan metode statistik non-parametrik melalui modul interaktif berbasis web menggunakan R Shiny.

2. memberikan contoh visualisasi data yang disesuaikan dengan tipe data dan metode pengujian statistik non parametrik.

2. Metodologi Penelitian

2.1 Tipe Penelitian

Penelitian yang dilakukan pada ini merupakan jenis penelitian deskriptif melalui pendekatan kualitatif serta kuantitatif. Pendekatan kualitatif pada penelitian deskriptif dilakukan untuk memberikan penggambaran sederhana terkait pengujian data dengan menggunakan metode statistik non-parametrik beserta contoh kasusnya menggunakan R Shiny. Pendekatan kuantitatif pada penelitian deskriptif dilakukan untuk memberikan penjelasan terkait metode pengujian, transformasi data, dan visualisasi data menggunakan R Shiny.

2.2 Teknik Pengambilan Data dari Sumber

- Dataset yang digunakan sebagai contoh dalam modul ini diambil dari berbagai sumber di internet dengan topik statistik non-parametrik khususnya di website universitas lain di fakultas statistika dan matematika seperti pada website [The University of Sheffield](#) serta website yang khusus memberikan dataset secara gratis seperti [StatGuru](#)
- Dataset yang digunkan diambil berdasarkan informasi dari data tersebut mengenai metode yang bisa digunakan untuk mengolah datanya

2.3 Teknik Analisis Data

Analisis dilakukan berdasarkan dataset yang telah ditransformasi kemudian dilakukan peninjauan berdasarkan hasil tes dan visualisasi data tersebut serta pembahasan terkait fitur yang terdapat pada

modul terkait.

3. Penetapan Variabel

variabel terdiri beberapa jenis diantaranya variabel bebas, variabel kontrol, variabel moderator, dan variabel terikat.

3.1.1 Jenis-Jenis Variabel

3.1.1.1 Jenis variabel Bebas (Independent Variabel)

Variabel bebas adalah variabel penelitian yang mempengaruhi variabel terikat. Variabel ini sering juga disebut sebagai variabel independen, stimulus, prediktor dan eksogen.

Contoh variabel bebas:

- Pengaruh motivasi belajar terhadap prestasi siswa di sekolah” motivasi belajar sebagai variabel bebas.
- Pengembangan fasilitas wisata dapat mempengaruhi variabel kepuasan pengunjung.
- Warna mobil bisa jadi variabel yang bisa mempengaruhi minat beli konsumen.

3.1.1.2 Variabel Terikat (Dependent Variabel)

Variabel terikat adalah variabel yang dipengaruhi oleh variabel bebas. Variabel ini juga disebut sebagai variabel independen atau variabel tergantung.

Contoh variabel terikat adalah “panjang batang tanaman yang dibedakan berdasarkan intensitas cahaya”.

3.1.1.3 Variabel Kontrol

Variabel kontrol adalah variabel penekan atau variabel yang dikendalikan oleh peneliti supaya dampak variabel bebas pada variabel terikat tidak terpengaruh oleh faktor di luar yang tidak diteliti.

Contoh variabel kontrol:

Hipotesis: Ada pengaruh metode pembelajaran terhadap penguasaan keterampilan akuntansi siswa kelas X SMK.

Variabel bebasnya adalah metode pembelajaran, misal metode praktek demonstrasi atau metode ceramah. Variabel kontrolnya, siswa dipilih berdasarkan tingkat kelas yang sama yaitu kelas X SMK. Dengan variabel kontrol tersebut, pengaruh metode pembelajaran dengan penguasaan keterampilan akuntansi siswa akan semakin akurat.

3.1.1.4 Variabel Moderator

Variabel moderator adalah variabel yang mempengaruhi hubungan antara variabel bebas dan terikat.

Hipotesis: Ada hubungan motivasi belajar dan prestasi siswa juga dipengaruhi oleh lingkungan belajarnya di sekolah.

Variabel bebas: motivasi belajar Contoh hubungan variabel moderator: iklim atau lingkungan belajar. Variabel terikat: prestasi siswa

Peran guru adalah membimbing dan menciptakan kondisi yang baik di sekolah. Apabila lingkungan belajar baik, maka hubungan motivasi belajar dan prestasi siswa akan meningkat.

3.1.1.1 Variabel Intervening

Variabel intervening adalah variabel yang mempengaruhi variabel bebas dan variabel terikat, namun tidak bisa diukur.

Contoh variabel intervening pada penelitian pengaruh pendapatan dengan umur harapan hidup di bawah ini adalah gaya hidup.

3.2 Penetapan Variabel Berdasarkan Dataset

3.2.1 Pengujian Binomial

- variabel bebas : nilai x (angka sukses), nilai n (percobaan), nilai p (hipotesis probabilitas sukses).
- variabel kontrol : toko yang sama.

- variabel moderator : alternative dan confidence level.
- variabel terikat : plot dan hasil tes.

3.2.2 Pengujian Wilcoxon

- variabel bebas : dataset sebelum, paired, alternative, hasil statistik deskriptif sebelum.
- variabel kontrol : suplemen penambah berat badan.
- variabel moderator : confidence level dan correct.
- variabel terikat : dataset sesudah, plot, hasil statistik deskriptif sesudah, dan hasil test.

3.2.3 Pengujian Mann-Whitney

- variabel bebas : hormon kortisol awal, paired, dan alternative.
- variabel kontrol : tugas yang diberikan kepada mahasiswa.
- variabel moderator : confidence level dan correct.
- variabel terikat : hormon kortisol akhir, plot, hasil statistik deskriptif, dan hasil test.

3.2.4 Pengujian McNemar

- variabel bebas : kategori awal partisipan dalam hal merokok.
- variabel kontrol : waktu selama 2 minggu, partisipan yang sama ketika direkrut, video bahaya merokok.
- variabel terikat : kategori akhir, hasil kontigensi, plot dan hasil tes.

3.2.5 Pengujian Spearman

- variabel bebas : temperature
- variabel kontrol : ice cream sales, pedagang yang sama.
- variabel moderator : confidence interval.
- variabel terikat ; correlation, statistika deskriptif, plot dan hasil tes

3.2.6 Pengujian Kruskal-Wallis

- variabel bebas : hormon kortisol.
- variabel kontrol : pasien dengan kategori a, b, dan c di rumah sakit yang sama.
- variabel terikat : statistika deskriptif, rank, plot dan hasil tes.

3.2.7 Pengujian Cochran

- variabel bebas : nilai siswa sebelum bimbel.
- variabel kontrol : tempat bimbel yang sama, partisipan yang sama, dan tujuan instansi yang sama.
- variabel terikat : hasil setelah bimbel, komposisi data, plot, dan hasil tes.

3.2.8 Pengujian Friedman

- variabel bebas : partisipan dari berbagai latar belakang.
- variabel kontrol : video a, b, c, dan d.
- variabel terikat : statistika deskriptif, rank, plot dan hasil tes.

Preprocessing

Preprocessing Data Uji Wilcoxon

```
1 # Wilcoxon test (Comparing two related conditions) data
```

```

2 sebelum <- c(25,27,20,21,18,19,20,22,24,25,24,27,23,25,22)
3 sesudah <- c(26,26,22,24,22,21,24,21,26,26,25,28,25,27,25)
4
5 data_wilcox = data.frame(sebelum,sesudah)
6
7 kable(
8   data_wilcox,
9   longtable = T,
10  booktabs = T,
11  linesep = c(""),
12  caption = 'Data untuk Uji Wilcoxon') %>%
13    kable_styling(latex_options = "striped")

```

Table 1: Data untuk Uji Wilcoxon

sebelum	sesudah
25	26
27	26
20	22
21	24
18	22
19	21
20	24
22	21
24	26
25	26
24	25
27	28
23	25
25	27
22	25

Preprocessing Data Uji Mann-Whitney

```

1 # Mann Whitney Signed Rank Test data
2 data_mw = read.csv("data_mw.csv")
3 kable(
4   data_mw,
5   longtable = T,
6   booktabs = T,
7   linesep = c(""),
8   caption = 'Data untuk Uji Mann-Whitney') %>%
9     kable_styling(latex_options = "striped")

```

Table 2: Data untuk Uji Mann-Whitney

cortisol_level_sebelum	cortisol_level_sesudah
3.97	33.26

4.67	7.37
14.00	18.83
17.64	62.70
8.86	8.92
4.21	15.85
11.37	16.36
8.03	20.74
3.99	58.79
9.75	10.27
2.45	1.30
2.65	8.52
9.29	13.21
13.04	13.04
16.43	12.59
8.01	20.82
NA	5.86
NA	19.11
NA	52.08
NA	15.69
NA	9.41
NA	17.13

Preprocessing Data Uji McNemar

```
1 # Mcnemar Test data
2 data_mc = read.csv("smoker.csv", sep = ',')
3 kable(
4   data_mc,
5   longtable = T,
6   booktabs = T,
7   linesep = c(""),
8   caption = 'Data untuk Uji McNemar') %>%
9   kable_styling(latex_options = "striped")
```

Table 3: Data untuk Uji McNemar

Before	After
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker

Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Non-Smoker
Non-Smoker	Smoker
Non-Smoker	Smoker
Non-Smoker	Smoker
Non-Smoker	Smoker
Non-Smoker	Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Non-Smoker
Smoker	Smoker
Smoker	Smoker
Smoker	Smoker
Smoker	Smoker
Smoker	Smoker
Smoker	Smoker
Smoker	Smoker

Preprocessing Data Uji Spearman

```

1 # Spearman Test data
2 data_sp = read.csv('spearman-correlation-data.csv', sep = ',')
3 kable(
4   data_sp,
5   longtable = T,
6   booktabs = T,
7   linesep = c(""),
8   caption = 'Data untuk Uji Spearman') %>%
9   kable_styling(latex_options = "striped")

```

Table 4: Data untuk Uji Spearman

ice_cream_sales	temperature
104	28
98	10
75	18
93	16
22	16
23	20
110	27
92	19
87	21
99	31
77	10
12	2
43	16
78	28
12	10
99	9
94	29
92	23
0	3
16	12

Preprocessing Data Uji Cochran

```

1 # Cochran Test data
2 data_cochran = read.csv('les.csv', sep = ',')
3 kable(
4   data_cochran,
5   longtable = T,
6   booktabs = T,
7   linesep = c(""),
8   caption = 'Data untuk Uji Cochran') %>%
9   kable_styling(latex_options = "striped")

```


Table 5: Data untuk Uji Cochran

Participant	Pre	Post
1	1	0
2	0	0
3	0	1
4	0	1
5	0	0
6	1	1
7	0	0
8	0	1
9	0	1
10	0	1

Preprocessing Data Uji Kruskal-Wallis

```

1 # kruskall-Wallis Test data
2 data_kw_base = read.csv('kruskal-wallis-test-data.csv', sep = ',')
3
4 value <- c(data_kw_base$category_a, data_kw_base$category_b, data_kw_base$category
   _c)
5 category <- c(rep(1,25), rep(2,25), rep(3,25))
6
7 data_kw <- data.frame(category, value)
8
9 kable(
10   data_kw_base,
11   longtable = T,
12   booktabs = T,
13   linesep = c(""),
14   caption = 'Data untuk Uji Kruskal-Wallis sebelum diolah') %>%
15   kable_styling(latex_options = "striped")

```

Table 6: Data untuk Uji Kruskal-Wallis sebelum diolah

category_a	category_b	category_c
33.26	3.97	19.05
7.37	4.67	9.45
18.83	14.00	3.45
62.70	17.64	0.32
7.10	8.86	5.32
8.92	4.21	2.92
15.85	11.37	2.33
16.36	8.03	5.62
20.74	3.99	3.98
58.79	9.75	19.03
10.27	2.45	0.33

1.30	2.65	15.33
22.70	9.29	3.22
8.52	13.04	3.96
4.21	16.43	4.09
13.21	8.01	9.65
13.04	2.61	4.34
12.59	10.61	1.11
10.48	11.34	4.33
20.82	1.54	2.22
5.86	18.49	5.43
19.11	16.87	0.98
52.08	12.93	3.88
15.69	4.99	2.06
9.41	4.32	1.03

```

1 kable(
2   data_kw,
3   longtable = T,
4   booktabs = T,
5   linesep = c(""),
6   caption = 'Data untuk Uji Kruskal-Wallis setelah diolah') %>%
7   kable_styling(latex_options = "striped")

```

Table 7: Data untuk Uji Kruskal-Wallis setelah diolah

category	value
1	33.26
1	7.37
1	18.83
1	62.70
1	7.10
1	8.92
1	15.85
1	16.36
1	20.74
1	58.79
1	10.27
1	1.30
1	22.70
1	8.52
1	4.21
1	13.21
1	13.04
1	12.59

1	10.48
1	20.82
1	5.86
1	19.11
1	52.08
1	15.69
1	9.41
2	3.97
2	4.67
2	14.00
2	17.64
2	8.86
2	4.21
2	11.37
2	8.03
2	3.99
2	9.75
2	2.45
2	2.65
2	9.29
2	13.04
2	16.43
2	8.01
2	2.61
2	10.61
2	11.34
2	1.54
2	18.49
2	16.87
2	12.93
2	4.99
2	4.32
3	19.05
3	9.45
3	3.45
3	0.32
3	5.32
3	2.92
3	2.33
3	5.62
3	3.98
3	19.03
3	0.33

3	15.33
3	3.22
3	3.96
3	4.09
3	9.65
3	4.34
3	1.11
3	4.33
3	2.22
3	5.43
3	0.98
3	3.88
3	2.06
3	1.03

Preprocessing Data Uji Friedman

```

1 # Friedman Test data
2
3 data_fried_base <- read.csv('Video_R.csv', sep = ',')
4
5 data_fried_base_table <- data.frame(data_fried_base$Person, data_fried_base$
  TotalAGen, data_fried_base$TotalBdoc, data_fried_base$TotalC0ld, data_fried_base
  $TotalDDEMO)
6
7 colnames(data_fried_base_table) = c('Person', 'Video A', 'Video B', 'Video C', '
  Video D')
8
9 video_score <- c(data_fried_base$TotalAGen, data_fried_base$TotalBdoc, data_fried_
  base$TotalC0l, data_fried_base$TotalDDEMO)
10 video_category <- c(rep(1,20), rep(2,20), rep(3,20), rep(4,20))
11
12 data_fried <- data.frame(data_fried_base$Person, video_category, video_score)
13
14 kable(
15   data_fried_base_table,
16   longtable = T,
17   booktabs = T,
18   linesep = c(""),
19   caption = 'Data untuk Uji Friedman sebelum diolah') %>%
20   kable_styling(latex_options = "striped")

```

Table 8: Data untuk Uji Friedman sebelum diolah

Person	Video A	Video B	Video C	Video D
1	25	23	13	22
2	23	23	20	22
3	20	17	14	23
4	24	24	23	25

5	25	22	18	23
6	24	22	23	25
7	25	22	9	21
8	24	23	6	21
9	25	21	11	23
10	24	23	5	25
11	24	23	15	21
12	22	24	14	19
13	24	23	14	19
14	25	25	17	18
15	24	22	14	17
16	25	23	13	20
17	25	22	16	21
18	25	21	11	22
19	25	23	19	25
20	24	23	18	25

```

1 kable(
2   data_fried,
3   longtable = T,
4   booktabs = T,
5   linesep = c(""),
6   caption = 'Data untuk Uji Friedman setelah diolah') %>%
7   kable_styling(latex_options = "striped")

```

Table 9: Data untuk Uji Friedman setelah diolah

data_fried_base.Person	video_category	video_score
1	1	25
2	1	23
3	1	20
4	1	24
5	1	25
6	1	24
7	1	25
8	1	24
9	1	25
10	1	24
11	1	24
12	1	22
13	1	24
14	1	25
15	1	24
16	1	25

17	1	25
18	1	25
19	1	25
20	1	24
1	2	23
2	2	23
3	2	17
4	2	24
5	2	22
6	2	22
7	2	22
8	2	23
9	2	21
10	2	23
11	2	23
12	2	24
13	2	23
14	2	25
15	2	22
16	2	23
17	2	22
18	2	21
19	2	23
20	2	23
1	3	13
2	3	20
3	3	14
4	3	23
5	3	18
6	3	23
7	3	9
8	3	6
9	3	11
10	3	5
11	3	15
12	3	14
13	3	14
14	3	17
15	3	14
16	3	13
17	3	16
18	3	11
19	3	19

20	3	18
1	4	22
2	4	22
3	4	23
4	4	25
5	4	23
6	4	25
7	4	21
8	4	21
9	4	23
10	4	25
11	4	21
12	4	19
13	4	19
14	4	18
15	4	17
16	4	20
17	4	21
18	4	22
19	4	25
20	4	25

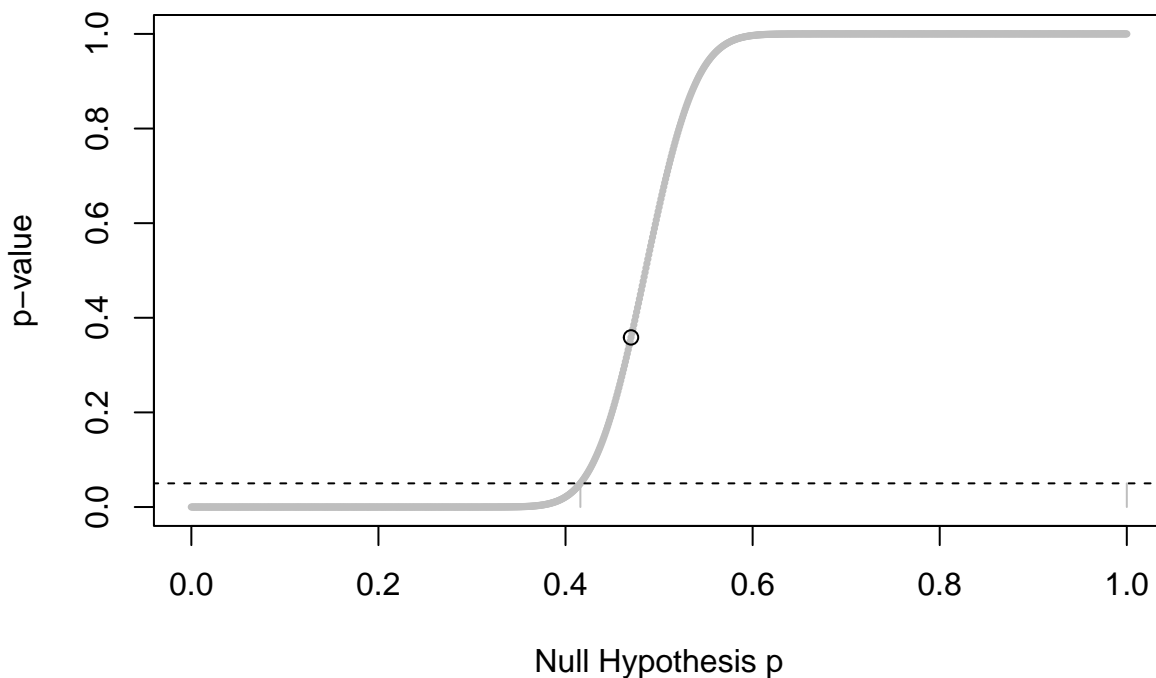
Pemodelan Statistika Deskriptif, Uji Non-Parametrik dan Visualisasi Data

Pemodelan Binomial

```

1 # binom.test(x, n, p = p_binom(), alternative = alternative_binom(), conf.level
  = ci_binom())
2
3 binom.exact(67, 137, p = 0.47, alternative = "greater", conf.level = 0.95,
  tsmethod = 'central', plot = TRUE)

```



```

1 ##
2 ##  Exact one-sided binomial test
3 ##
4 ## data:  67 and 137
5 ## number of successes = 67, number of trials = 137, p-value = 0.3585
6 ## alternative hypothesis: true probability of success is greater than 0.47
7 ## 95 percent confidence interval:
8 ##  0.4158639 1.0000000
9 ## sample estimates:
10 ## probability of success
11 ##          0.4890511

```

Pemodelan Wilcoxon

```

1 result <- wilcox.test(sebelum, sesudah, paired = TRUE, correct = TRUE, conf.level
  = 0.95, alternative = 'two.sided')

```

```

1 ## Warning in wilcox.test.default(sebelum, sesudah, paired = TRUE, correct =
2 ## TRUE, : cannot compute exact p-value with ties

```

```

1 result

```

```

1 ##
2 ##  Wilcoxon signed rank test with continuity correction
3 ##
4 ## data:  sebelum and sesudah
5 ## V = 7, p-value = 0.002556
6 ## alternative hypothesis: true location shift is not equal to 0

```

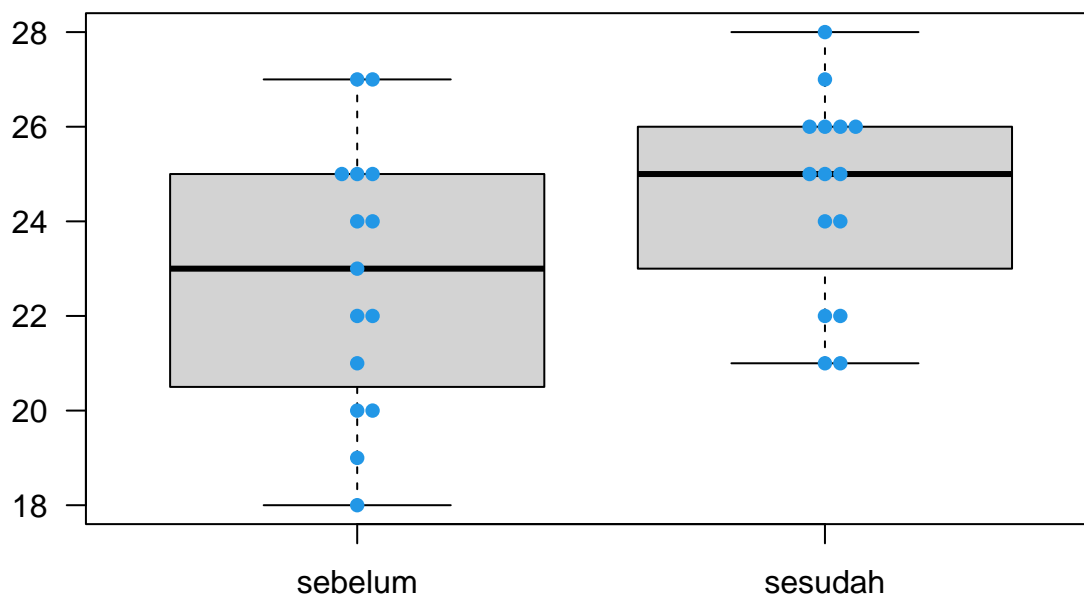


```
1 summary_data_wilcox <- summary(data_wilcox)
2 summary_data_wilcox
```

```
1 ##      sebelum      sesudah
2 ## Min.   :18.0    Min.   :21.00
3 ## 1st Qu.:20.5    1st Qu.:23.00
4 ## Median :23.0    Median :25.00
5 ## Mean   :22.8    Mean   :24.53
6 ## 3rd Qu.:25.0    3rd Qu.:26.00
7 ## Max.   :27.0    Max.   :28.00
```

```
1 boxplot(data_wilcox, las=1, main = 'Perbandingan berat badan anak')
2 beeswarm(data_wilcox, col = 4, pch = 16, vert = TRUE, add = TRUE)
```

Perbandingan berat badan anak



Pemodelan Mann-Whitney

```
1 result <- wilcox.test(data_mw$cortisol_level_sebelum, data_mw$cortisol_level_
  sesudah, paired = TRUE, correct = TRUE, conf.level = 0.95, alternative = 'two.
  sided')
```

```
1 ## Warning in wilcox.test.default(data_mw$cortisol_level_sebelum,
2 ## data_mw$cortisol_level_sesudah, : cannot compute exact p-value with zeroes
```

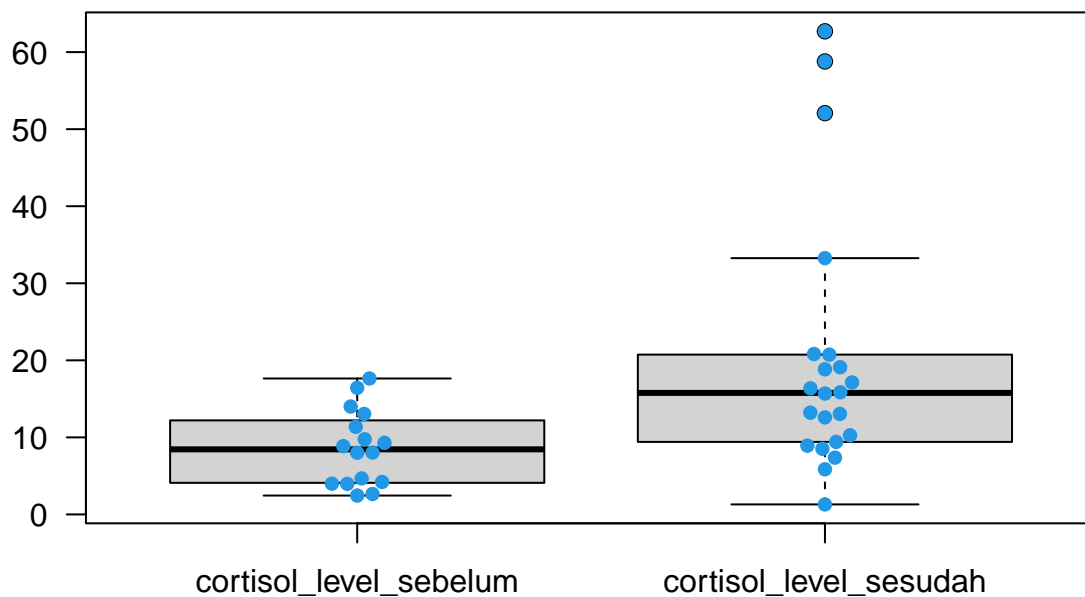
```
1 result
```

```
1 ##
2 ## Wilcoxon signed rank test with continuity correction
3 ##
4 ## data: data_mw$cortisol_level_sebelum and data_mw$cortisol_level_sesudah
```

```
5 ## V = 8, p-value = 0.003445
6 ## alternative hypothesis: true location shift is not equal to 0
```

```
1 boxplot(data_mw, las=1, main = 'Perbandingan hormon kortisol (mcg/dL).')
2 beeswarm(data_mw, col = 4, pch = 16, vert = TRUE, add = TRUE)
```

Perbandingan hormon kortisol (mcg/dL).



Pemodelan McNemar

```
1 x <- table(data_mc)
2 x <- addmargins(x)
3
4 smoker <- table(data_mc)
5 res1 <- mcnemar.test(smoker)
6 res2 <- mcnemar.exact(smoker)
7
8 McNemarChi <- paste("McNemar's chi-squared = ", round(res1[[1]][[1]],3), ", ", " ", "
   df = ", res1[[2]][[1]], sep = "")
9 cat(sprintf(McNemarChi), "\n")
```

```
1 ## McNemar's chi-squared = 4.762, df = 1
```

```
1 print(res2)
```

```
1 ##
2 ## Exact McNemar test (with central confidence intervals)
3 ##
4 ## data:  smoker
5 ## b = 5, c = 16, p-value = 0.0266
6 ## alternative hypothesis: true odds ratio is not equal to 1
7 ## 95 percent confidence interval:
```

```

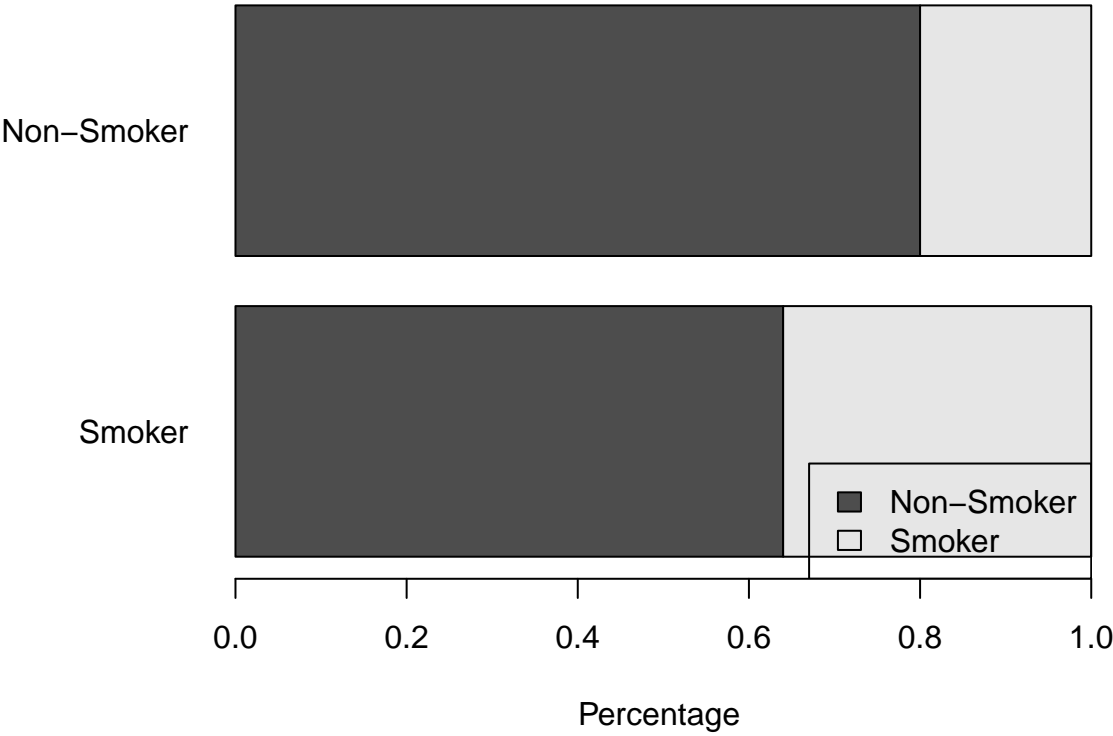
8  ##  0.08953336 0.89271998
9  ## sample estimates:
10 ## odds ratio
11 ##      0.3125

```

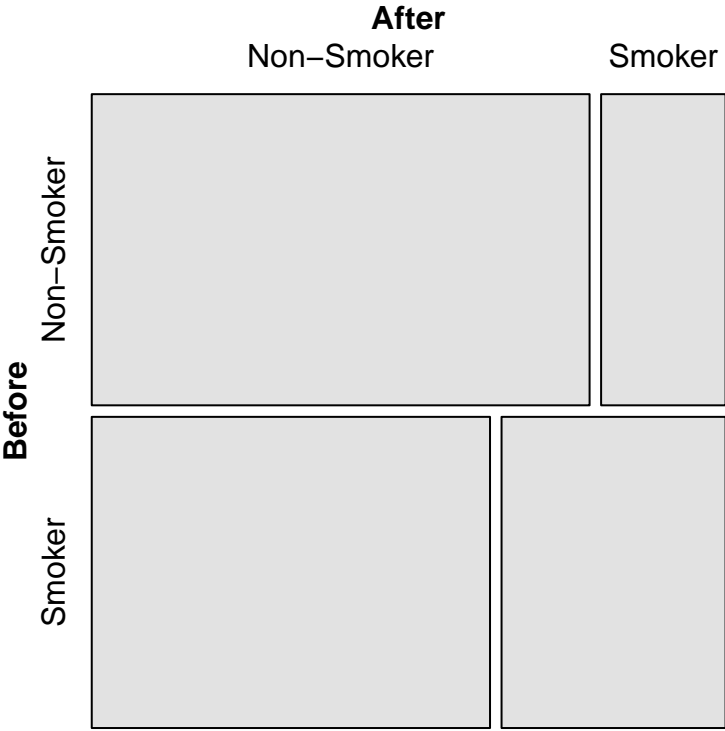
```

1  makePlot <- function() {
2    x <- table(data_mc)
3
4    levI <- nrow(x)
5    levJ <- ncol(x)
6    matrix_invers <- as.vector(t(x))
7
8    total <- c()
9    total_rep <- c()
10   for(i in 1:levI)
11   {
12     ds <- c()
13     for(j in 1:levJ)
14     {
15       ds <- c(ds, matrix_invers[(i-1)*levJ+j])
16     }
17     total <- c(total, sum(ds))
18     total_rep <- c(total_rep, rep(sum(ds), levJ))
19   }
20   percentage <- matrix_invers/total_rep
21
22   a <- c()
23   for(i in levI:1)
24   {
25     for(j in 1:levJ)
26     {
27       a <- c(a, percentage[(i-1)*levJ+j] )
28     }
29   }
30
31   b <- matrix(c(a), nc=levJ, by=1)
32   rownames(b) <- rev(rownames(x))
33   colnames(b) <- colnames(x)
34
35   par(mar=c(5,6,2,4))
36   barplot(t(b), hor=1, las=1, xlab="Percentage", col=gray.colors(ncol(x)))
37   legend("bottomright", legend=colnames(b), fill=gray.colors(ncol(x)))
38 }
39
40 makemPlot <- function(){
41   x <- table(data_mc)
42   mosaic(x, gp = shading_max, legend=FALSE)
43 }
44
45 makePlot()

```



```
1 makemPlot()
```



Pemodelan Spearman

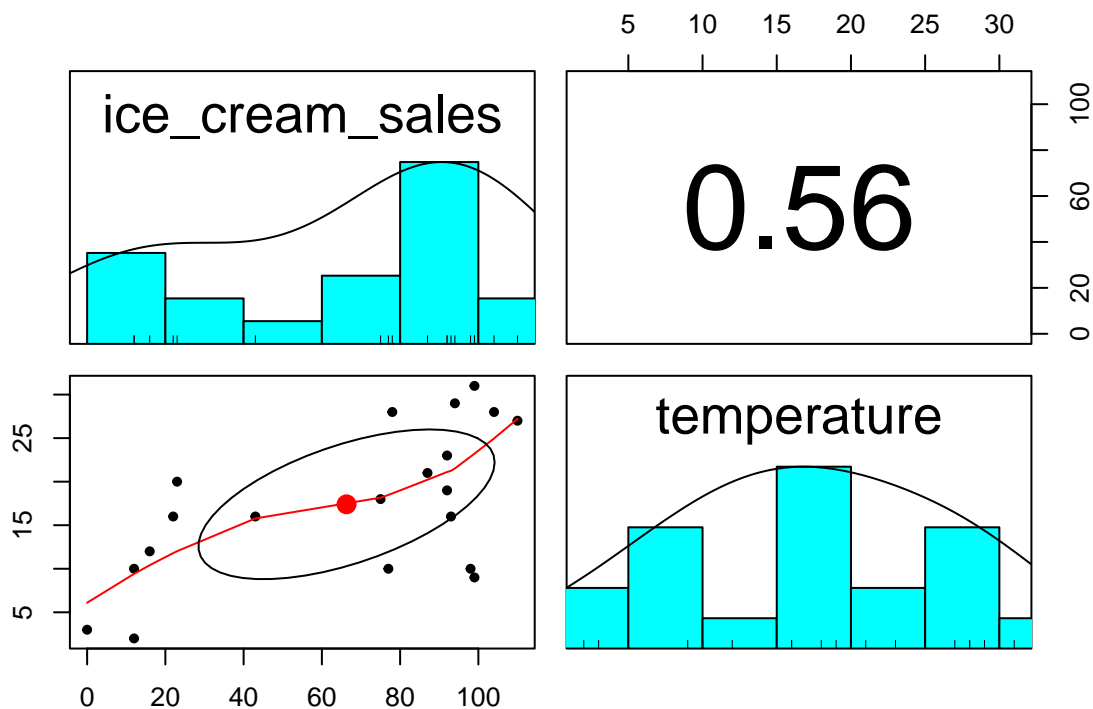
```
1 data_spearman = as.matrix(data_sp)
2 describe(data_spearman)
```

```
1 ##           vars  n mean    sd median trimmed  mad min max range  skew
2 ## ice_cream_sales    1 20 66.3 37.86   82.5   68.75 24.46   0 110   110 -0.55
3 ## temperature        2 20 17.4  8.60   17.0   17.69 10.38   2  31    29 -0.08
4 ##
5 ##           kurtosis    se
6 ## ice_cream_sales   -1.49 8.47
7 ## temperature      -1.16 1.92
```

```
1 x <- as.matrix(data_sp)
2 round(cor(cbind(x), method = 'spearman', use = "pairwise.complete.obs"),3)
```

```
1 ##           ice_cream_sales temperature
2 ## ice_cream_sales           1.000         0.561
3 ## temperature              0.561         1.000
```

```
1 x <- as.matrix(data_sp)
2 pairs.panels(x, method = 'spearman')
```



Pemodelan Kruskal-Wallis

```
1 describeBy(data_kw[,2], data_kw[,1])
```

```
1 ##
2 ## Descriptive statistics by group
3 ## group: 1
4 ##           vars  n mean    sd median trimmed  mad min  max range skew kurtosis  se
5 ## X1          1 25 18.77 16.31  13.21   16.3  8.66 1.3 62.7  61.4 1.57    1.42 3.26
```

```

6  ## -----
7  ## group: 2
8  ##      vars  n mean    sd median trimmed  mad   min    max range skew kurtosis   se
9  ## X1      1 25 8.88 5.26   8.86   8.66 6.73 1.54 18.49 16.95  0.3    -1.26 1.05
10 ## -----
11 ## group: 3
12 ##      vars  n mean    sd median trimmed  mad   min    max range skew kurtosis   se
13 ## X1      1 25 5.34 5.28   3.96   4.51 2.46 0.32 19.05 18.73 1.54     1.31 1.06

```

```

1 KW.ranking <- function() {
2   ranked <- rank(data_kw[,2])
3   data <- data.frame(data_kw[,1], ranked)
4   n <- round(tapply(data[,2], data[,1], length),2)
5   m <- round(tapply(data[,2], data[,1], mean),2)
6   t <- round(tapply(data[,2], data[,1], sum),2)
7   ranks <- data.frame(n, m, t)
8   colnames(ranks) <- c("n", "Rank Mean", "Rank Sum")
9   print(ranks)
10 }
11 KW.ranking()

```

```

1  ##      n Rank Mean Rank Sum
2  ## 1 25      52.72     1318
3  ## 2 25      37.80      945
4  ## 3 25      23.48      587

```

```

1 result <- kruskal.test(data_kw[,2] ~ data_kw[,1])
2 print(result)

```

```

1  ##
2  ##  Kruskal-Wallis rank sum test
3  ##
4  ## data:  data_kw[, 2] by data_kw[, 1]
5  ## Kruskal-Wallis chi-squared = 22.503, df = 2, p-value = 1.299e-05

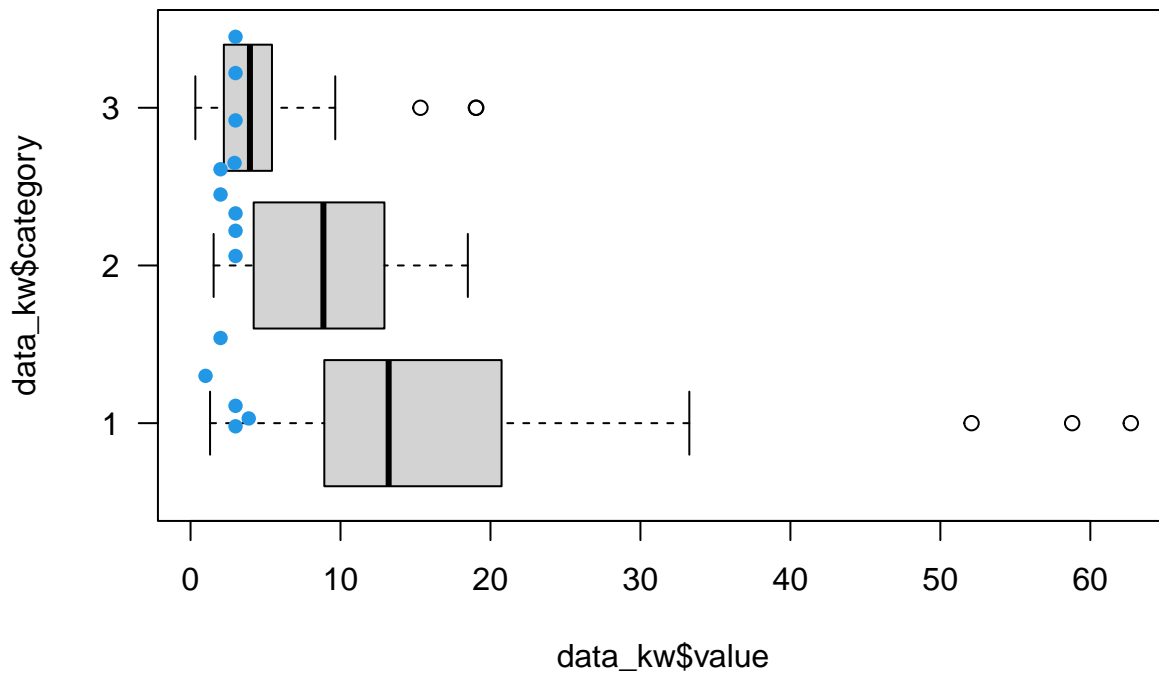
```

```

1 dat <- data_kw
2 boxplot(data_kw$value ~ data_kw$category, las=1, horizontal = TRUE, main = '
  Perbandingan hormon kortisol pasien (mcg/dL).')
3 beeswarm(dat[,2] ~ dat[,1], col = 4, pch = 16, vert = TRUE, add = TRUE)

```

Perbandingan hormon kortisol pasien (mcg/dL).



Pemodelan Cochran

```
1 data.cochran <- function() {
2   dat <- data_cochran
3   dat[,1] <- factor(dat[,1])
4   data.long <- melt(dat, idvars=dat[,1])
5   x <- t(table(data.long$variable, data.long$value))
6   x <- addmargins(x)
7   print(x)
8 }
9
10 data.cochran()
```

```
1 ## Using Participant as id variables
```

```
1 ##
2 ##      Pre Post Sum
3 ##      0      8   4  12
4 ##      1      2   6   8
5 ##      Sum  10  10  20
```

```
1 cochran.test <- function() {
2   dat <- data_cochran
3
4   dat[,1] <- factor(dat[,1])
5   data.long <- melt(dat, idvars=dat[,1])
6   q <- symmetry_test(data.long[,3] ~ factor(data.long[,2]) | factor(data.
7     long[,1]), data=data.long, teststat="quad")
8   CochranQChi <- paste("Cochran's Q chi-squared = ", round(
```

```

    q@statistic@teststatistic,3), ", ", "df = ", q@statistic@df, sep = "")
9   cat(sprintf(CochranQChi), "\n")
10
11   P.CochranQChi <- paste("p-value = ", pvalue(q), sep = "")
12   cat(sprintf(P.CochranQChi), "\n", "\n")
13
14
15   cat("Effect size for Cochran's Q test:", "\n")
16   eta.squared.q <- q@statistic@teststatistic / (nrow(dat) * ((ncol(dat)-1)
    -1))
17   ESQ <- paste("Eta-squared Q = ", round(eta.squared.q,3), sep = "")
18   cat(sprintf(ESQ), "\n", "\n", "\n")
19 }
20
21 cochran.test()

```

```
1 ## Using Participant as id variables
```

```

1 ## Cochran's Q chi-squared = 2.667, df = 1
2 ## p-value = 0.102470434859749
3 ##
4 ## Effect size for Cochran's Q test:
5 ## Eta-squared Q = 0.267
6 ##
7 ##

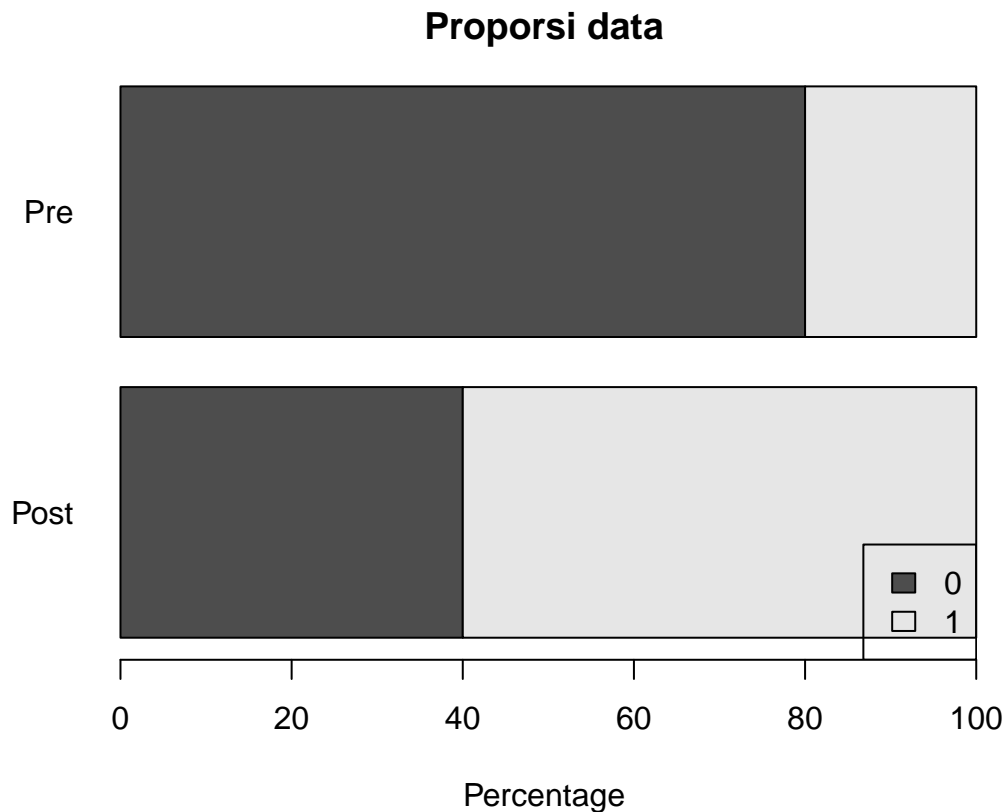
```

```

1 makePlot.ch <- function(){
2
3   dat <- data_cochran
4
5   dat[,1] <- factor(dat[,1])
6   data.long <- melt(dat, idvars=dat[,1])
7   x <- t(table(data.long$variable, data.long$value))
8   n <- nrow(dat)
9   prp <- round(((x/n)*100), 1)
10  prp.rev <- apply(prp, 1, rev)
11
12  par(mar=c(5,6,2,4))
13  barplot(t(prp.rev), hor=1, las=1, xlab="Percentage", main = 'Proporsi data')
14  legend("bottomright", legend=rownames(x), fill=gray.colors(nrow(x)))
15 }
16
17 makePlot.ch()

```

```
1 ## Using Participant as id variables
```

Pemodelan Friedman

```
1 describeBy(data_fried[,3], data_fried[,2])
```

```
1 ##
2 ## Descriptive statistics by group
3 ## group: 1
4 ##   vars  n mean   sd median trimmed  mad min max range  skew kurtosis   se
5 ## X1    1 20 24.1 1.25    24   24.38 1.48  20  25    5 -1.86    3.29 0.28
6 ## -----
7 ## group: 2
8 ##   vars  n mean   sd median trimmed  mad min max range  skew kurtosis   se
9 ## X1    1 20 22.45 1.61    23   22.62 1.48  17  25    8 -1.73    4.13 0.36
10 ## -----
11 ## group: 3
12 ##   vars  n mean   sd median trimmed  mad min max range  skew kurtosis   se
13 ## X1    1 20 14.65 4.87    14   14.75 4.45   5  23   18 -0.15   -0.61 1.09
14 ## -----
15 ## group: 4
16 ##   vars  n mean   sd median trimmed  mad min max range  skew kurtosis   se
17 ## X1    1 20 21.85 2.46    22    22  2.22  17  25    8 -0.22   -1.05 0.55
```

```
1 fried.ranking <- function() {
2   ranked <- rank(data_fried[,3])
3   data <- data.frame(data_fried[,2], ranked)
4   n <- round(tapply(data[,2], data[,1], length),2)
5   m <- round(tapply(data[,2], data[,1], mean),2)
6   t <- round(tapply(data[,2], data[,1], sum),2)
7   ranks <- data.frame(n, m, t)
8   colnames(ranks) <- c("n", "Rank Mean", "Rank Sum")
9 }
```

```

9   print(ranks)
10  }
11
12  fried.ranking()

```

```

1  ##      n Rank Mean Rank Sum
2  ## 1 20      62.45      1249
3  ## 2 20      43.90       878
4  ## 3 20      14.15       283
5  ## 4 20      41.50       830

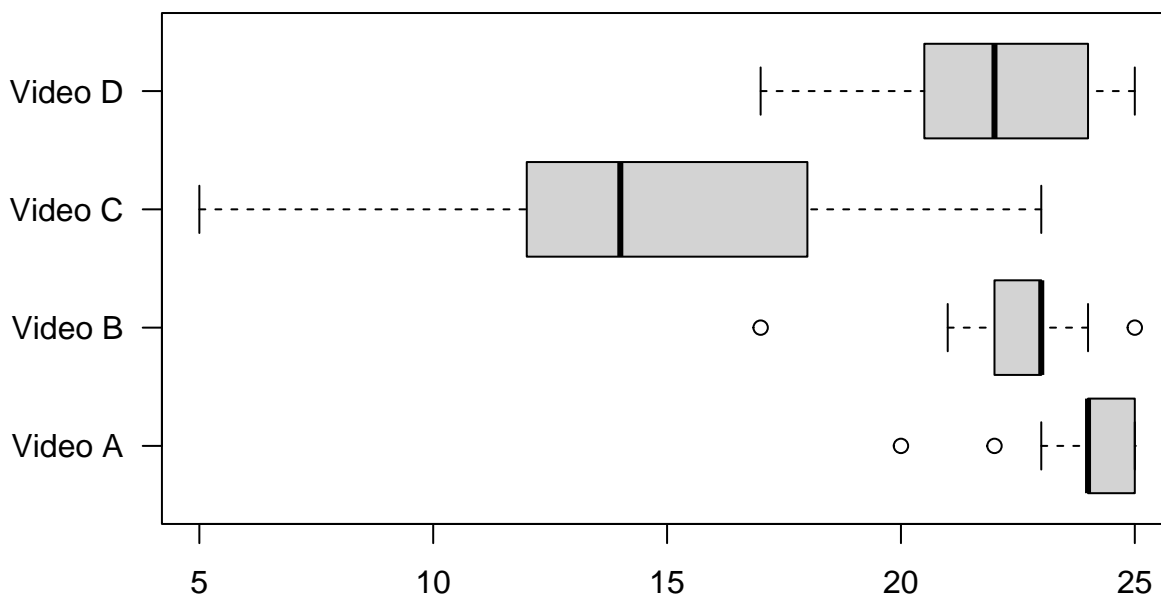
```

```

1  fried.test <- function() {
2    data_video_score = as.matrix(data_fried)
3    result <- friedman.test(data_video_score)
4    print(result)
5  }
6
7  friedman.boxPlot <- function(){
8    dat <- data_fried_base_table[,-1]
9    boxplot(dat, las=1, horizontal = TRUE, main = 'Perbandingan Score Tiap Video')
10   beeswarm(dat, col = 4, pch = 16, vert = TRUE, add = TRUE)
11 }
12
13 friedman.boxPlot()

```

Perbandingan Score Tiap Video



Pemodelan Aplikasi

SDT PENS App Beta

Cover

Binomial Test

Wilcoxon Test

Mann-Whitney Test

McNemar Test

Spearman Correlation Test

Kruskal-Wallis Test

Cochran Test

Friedman Test

Binomial Test for Statistics Non-Parametric

Description :
Uji Binomial adalah uji non parametrik yang digunakan untuk menguji hipotesis bila dalam populasi terdiri atas dua kelompok kelas dengan tipe data binomial dan ukuran sampelnya kecil. Uji Binomial menggantikan uji statistik t jika asumsi n kecil dan populasi normal sebagai syarat uji t tidak dipenuhi.

Sample Case

Terdapat berbagai jenis kayu untuk pembuatan furnitur rumah tangga, seperti Kayu Jati, Kayu Mahogani, Kayu Trembesi, Kayu Sungkai, Kayu Mindi. Sebelum pandemi datang Kayu Jati menjadi jenis kayu yang paling laris dengan total penjualan sebesar 47% dari keseluruhan penjualan, tetapi Semenjak pandemi terjadi pemesanan jenis Kayu Jati disinyalir mengalami penurunan karena memang Kayu Jati merupakan jenis kayu yang paling mahal. Oleh karena itu, dilakukan penelitian dari toai penjualan setelah pandemi dan didapatkan data bahwa dari 137 juta log kayu yang terjual, pemesanan dengan jenis Kayu Jati sebesar 67 juta log.

Kasus ini dapat dibawa ke persoalan binomial yang mana p merupakan proporsi Kayu Jati yang dimpesan. Sehingga jenis kayu selain itu berarti bukan Kayu Jati.

Options

Nilai x

67

Nilai n

137

Nilai p

0.47

Conf.level

0.95

Alternative

Greater

Result and Plot

```

Exact binomial test

data: x and n
number of successes = 67, number of trials = 137, p-value = 0.3585
alternative hypothesis: true probability of success is greater than 0.47
95 percent confidence interval:
 0.4158639 1.0000000
sample estimates:
probability of success
 0.4890511
          
```

Wilcoxon Test for Statistics Non-Parametric

Description :
Uji wilcoxon merupakan uji non parametrik yang digunakan untuk mengukur ada tidaknya perbedaan nilai rata-rata dua kelompok sampel yang saling berpasangan (dependen).

Sample Case

Sebuah perusahaan Pharmasi sedang mengembangkan suplemen penambahan berat badan pada anak-anak. Perusahaan ingin mengetahui khasiat suplemen tersebut sebelum dipasarkan secara komersial. Untuk itu perusahaan mencoba obat tersebut secara kontinu terhadap 15 orang siswa sekolah dasar yang sudah diukur terlebih dahulu berat badannya. Setelah 3 bulan kemudian siswa-siswa tersebut ditimbang berat badannya lagi untuk mengetahui apakah ada peningkatan berat badannya yang nyata.

Dataset

Show 10 entries

Search:

	sebelum	sesudah
1	25	26
2	27	26
3	20	22
4	21	24
5	18	22
6	19	21

Plot

Perbandingan berat badan anak

SDT PENS App Beta

Cover

Binomial Test

Wilcoxon Test

Mann-Whitney Test

McNemar Test

Spearman Correlation Test

Kruskal-Wallis Test

Cochran Test

Friedman Test

Mann-Whitney Test for Statistics Non-Parametric

Description :
Uji Mann-Whitney merupakan uji non parametrik yang digunakan untuk mengetahui perbedaan median 2 kelompok bebas dengan skala data pengukuran adalah ordinal atau interval maupun rasio, namun data tersebut tidak berdistribusi normal. Sintaksis uji tersebut di R adalah sama dengan uji Wilcoxon.

Sample Case

Peneliti melakukan eksperimen kepada beberapa mahasiswa untuk melihat total hormon kortisol (hormon pengendali stres) yang dihasilkan saat sebelum dan sesudah diberi tugas besar dengan deadline yang singkat

Dataset

Show 10 entries

Search:

	cortisol_level_sebelum	cortisol_level_sesudah
1	3.97	
2	4.67	
3	14	
4	17.64	
5	8.86	
6	4.21	
7	11.37	
8	8.03	
9	3.99	
10	9.75	

Showing 1 to 10 of 22 entries

Previous

1

2

3

Next

Plot

Options

Paired: TRUE

Correct: TRUE

Alternative: Two Sided

Conf.level:

Result and Descriptive Statistics

Result Mann-Whitney Test

```

Wilcoxon signed rank test with continuity correction

data: data_mv$cortisol_level_sebelum and data_mv$cortisol_level_sesudah
V = 8, p-value = 0.003445
alternative hypothesis: true location shift is not equal to 0

```

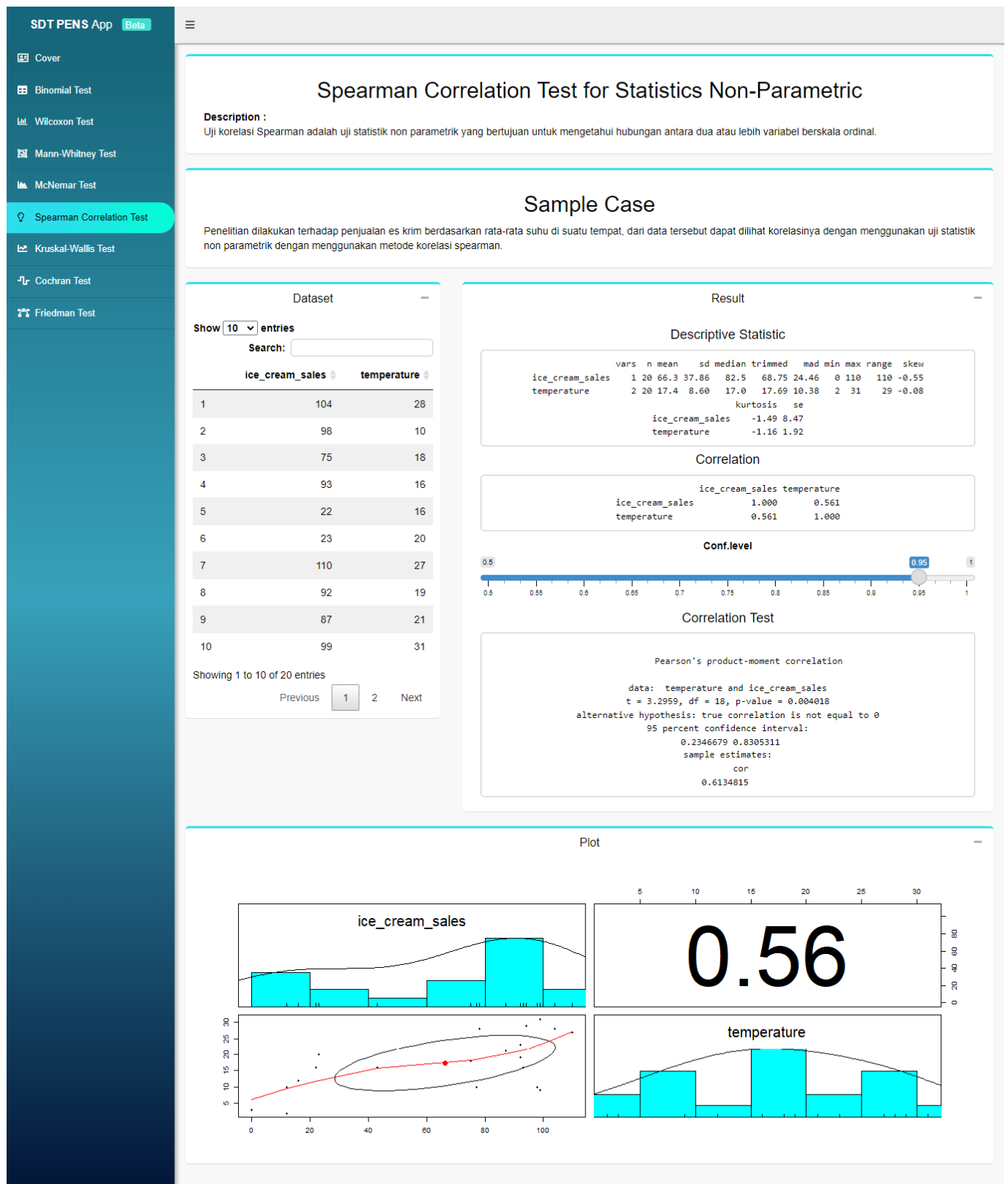
Descriptive Statistics

cortisol_level_sebelum	cortisol_level_sesudah
Min. : 2.450	Min. : 1.300
1st Qu.: 4.155	1st Qu.: 9.625
Median : 8.445	Median : 15.770
Mean : 8.648	Mean : 20.084
3rd Qu.: 11.787	3rd Qu.: 20.332
Max. : 17.640	Max. : 62.700
NA's : 6	

Muhammad Dzalhaqi (3321600023)

28





SDT PENS App Beta

- Cover
- Binomial Test
- Wilcoxon Test
- Mann-Whitney Test
- McNemar Test
- Spearman Correlation Test
- Kruskal-Wallis Test**
- Cochran Test
- Friedman Test

Kruskal-Wallis Test for Statistics Non-Parametric

Description :
 Uji kruskal Wallis adalah uji statistik non parametrik yang dapat digunakan untuk menguji apakah ada perbedaan yang signifikan antara kelompok variabel. Karena untuk melihat perbedaan yang signifikan antar kelompok, uji ini jelas digunakan untuk melihat perbandingan lebih dari 2 kelompok populasi dengan data berbentuk ranking. Uji ini juga disebut sebagai uji kruskal-wallis H, atau H-test, yang merupakan perluasan uji 2 sampel wilcoxon untuk $k > 2$ sampel. Uji Kruskal Wallis merupakan uji alternatif untuk uji F dan uji one way Anova (Analysis of Variance) untuk pengujian kesamaan beberapa nilai tengah dan analisis ragam yang dapat kita gunakan jika asumsi kenormalan tidak terpenuhi atau berdistribusi bebas.

Sample Case

Penelitian dilakukan terhadap 75 pasien dengan 3 kategori di rumah sakit dengan mengukur tingkat hormon kortisol yang dihasilkan. pasien dengan kategori A menandakan bahwa pasien memiliki sindrom cushing, pasien dengan kategori B menandakan bahwa pasien memiliki penyakit jantung koroner, pasien dengan kategori C menandakan bahwa pasien dengan pengontrolan hormon kortisol

Dataset

Show 10 entries

Search:

	category_a	category_b	category_c
1	33.26	3.97	1
2	7.37	4.67	
3	18.83	14	
4	62.7	17.64	
5	7.1	8.86	
6	8.92	4.21	
7	15.85	11.37	
8	16.36	8.03	
9	20.74	3.99	
10	58.79	9.75	1

Showing 1 to 10 of 25 entries

Previous 1 2 3 Next

Descriptive Statistics and Result

Descriptive Statistics

Descriptive statistics by group

group: 1													
vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se	
X1	1	25	18.77	16.31	13.21	16.3	8.66	1.3	62.7	61.4	1.57	1.42	3.26

group: 2													
vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se	
X1	1	25	8.88	5.26	8.86	8.66	6.73	1.54	18.49	16.95	0.3	-1.26	1.05

group: 3													
vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se	
X1	1	25	5.34	5.28	3.96	4.51	2.46	0.32	19.05	18.73	1.54	1.31	1.06

Rank

	n	Rank	Mean	Rank Sum
1	25	52.72	1318	
2	25	37.80	945	
3	25	23.48	587	

Result

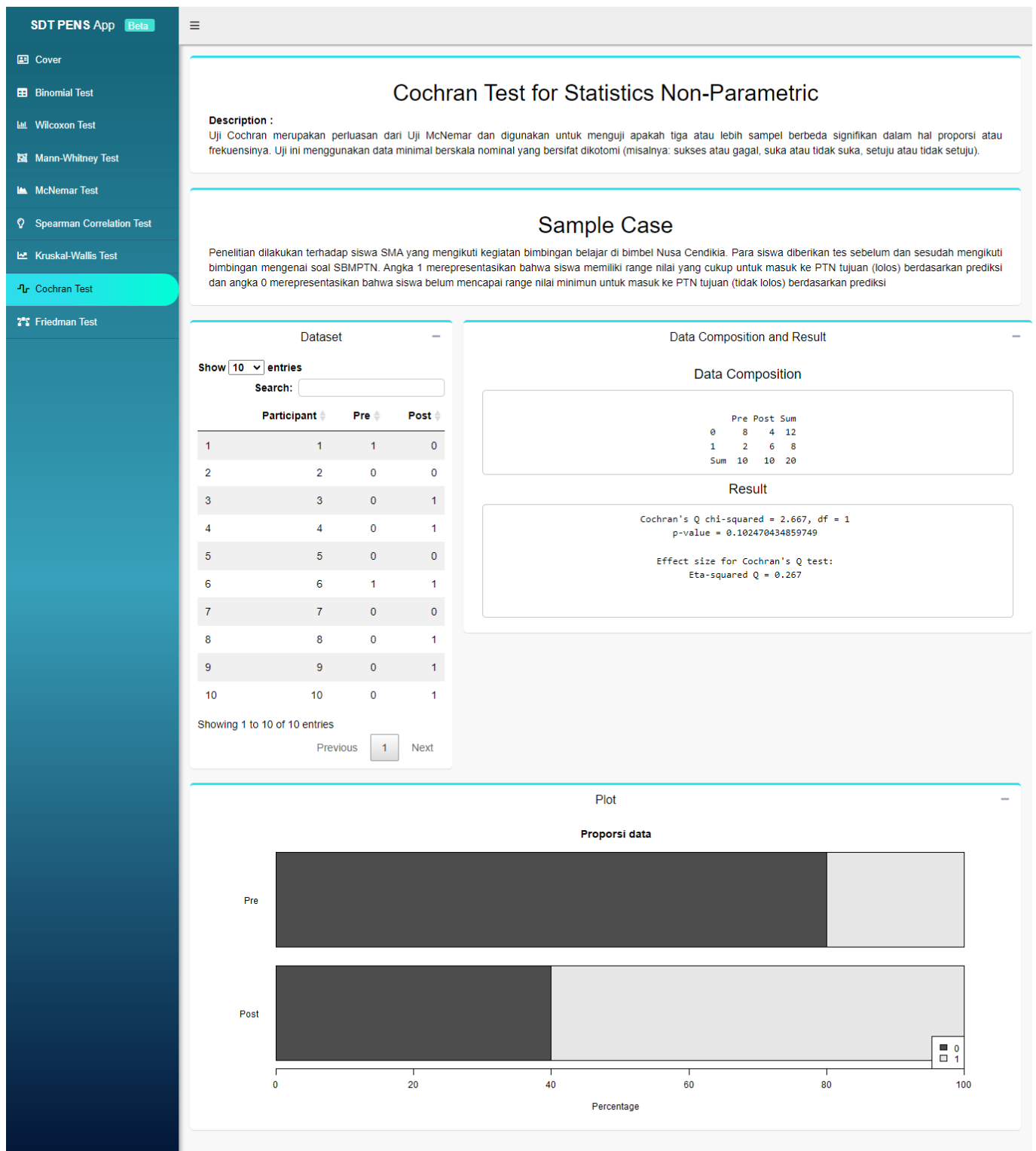
Kruskal-Wallis rank sum test

data: data_kw[, 2] by data_kw[, 1]

Kruskal-Wallis chi-squared = 22.503, df = 2, p-value = 1.299e-05

Plot

Perbandingan kategori pasien berdasarkan jumlah hormon kortisol dengan satuan microgram per deciliter (mcg/dL).



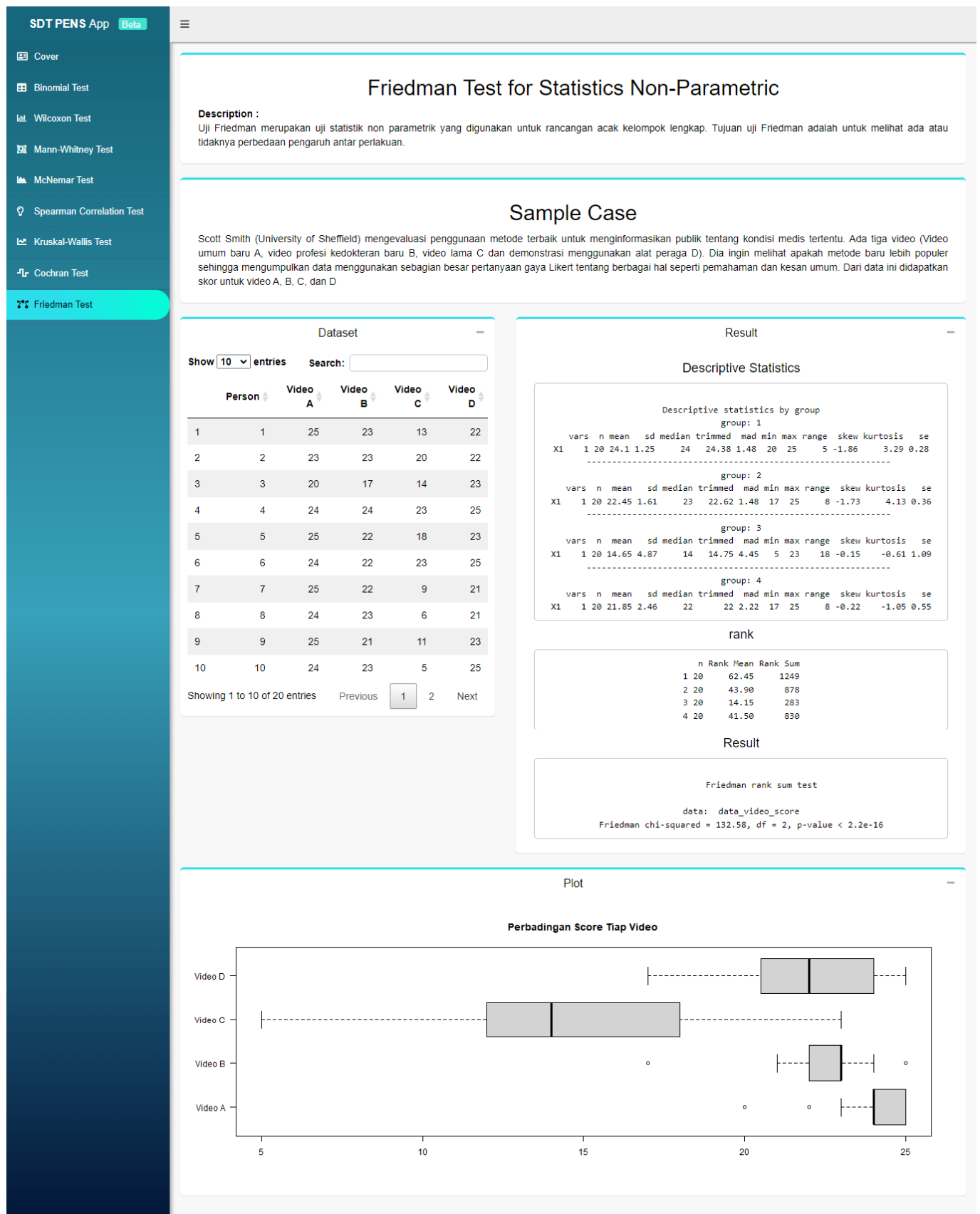




Figure 1: Tab Cover

Note

Setiap tab pada aplikasi ini memiliki fitur yang serupa, tetapi tidak sama persis. Secara general fitur yang ada pada setiap tab meliputi judul beserta deskripsi, contoh kasus, tampilan contoh dataset, box yang terdiri grup hasil perhitungan statistika, option (pilihan parameter) jika ada, dan Plot. setiap fungsi untuk menguji data menggunakan beberapa metode dalam statistika non-parametrik yang ada memiliki parameter yang berbeda pula. Begitu juga dengan hasil tes, bagian tersebut bergantung pada paramter yang yang dipilih/diinputkan oleh pengguna jikalau memang ada pilihan/inputan.

Evaluasi

Evaluasi dari Kode Program

1. Kode belum efektif dan efisien
2. Kode masih terlalu sulit untuk dipahami

Evaluasi dari Fitur Aplikasi

1. Fitur dan desain aplikasi masih sederhana
2. Beberapa contoh kasus tidak terlalu nyata

Kesimpulan

Seorang *data scientist* haruslah memiliki kecakapan dalam *programming* dan ilmu statistika sebab dua hal itu merupakan yang paling penting untuk mengembangkan ide dan membuat inovasi. Dalam hal ini diperlukan

latihan secara konsisten untuk melatih keterampilan data mengolah data. Dibuatnya aplikasi ini adalah salah tujuan untuk mengasah keterampilan serta pengetahuan penulis sebagai orang yang mempelajari ilmu data khususnya pada topik statistik non-parametrik.

Statistik non-paramterik adalah bagian ilmu statistika yang mempelajari cara mengolah data ketika sampel terlalu sedikit dan kita tidak memiliki cukup bukti untuk mempercayai bahwa data berdistribusi normal kecuali jika memang data tersebut diketahui berdistribusi kontinyu. Dengan cara ini kita bisa melakukan pengujian terhadap data dengan beberapa metode Statistik Non-Parametrik seperti Uji Binomial, Uji Wilcoxon, Uji Mann-Whitney, Uji McNemar, Uji Spearman, Uji Kruskall-Wallis, Uji Cochran, dan Uji Friedman. Setiap Metode pengujian memiliki fungsinya tersendiri dalam menguji datanya untuk menghasil kesimpulan yang dibutuhkan.

Daftar Pustaka

Shiny. <https://shiny.rstudio.com/>. Accessed 1 Dec 2021.

<https://mastering-shiny.org/>

Probability & Statistics for Engineers & Scientists (9th Edition)