

Analisis Preferensi Genre Game Berdasarkan Jenis Kelamin Mahasiswa UMN Angkatan 2021

¹*Daffa Raelby*, ²*Jovansha Cancerio*, ³*Muhammad Zaidan Fiqri*

^{1, 2, 3}Informatika, Universitas Multimedia Nusantara, Tangerang, Indonesia
daffa.raelby@student.umn.ac.id, jovansha.cancerio@student.umn.ac.id,
muhammad.zaidan@student.umn.ac.id

Abstrak - Penelitian ini dibuat dengan tujuan untuk mengetahui Genre game yang paling banyak digemari berdasarkan kategori. Melalui penelitian ini, didapatkanlah 89 responden yang sampelnya diambil dari mahasiswa universitas multimedia nusantara angkatan 2021. Penelitian ini menggunakan data primer dengan pengumpulan data dengan membuat kuesioner menggunakan layanan *google form* yang berisikan Pertanyaan *Numerik* dan *Kategorik* yang harus dijawab oleh Responden. Penelitian ini menghasilkan fakta-fakta sebagai berikut: (1) Rata-Rata waktu bermain game Mahasiswa Universitas Multimedia Nusantara adalah 3 jam lebih. (2) Game dengan Genre Action merupakan genre yang paling banyak dimainkan oleh Mahasiswa Universitas Multimedia Nusantara. (3) Rata-Rata waktu Durasi Rehat dari bermain game adalah 8 jam lebih.

I. PENDAHULUAN

A. latar belakang

Bermain game merupakan salah satu cara untuk melepas stress setelah beraktifitas seharian. Terdapat berbagai jenis game yang berbeda-beda yang terbagi melalui genre, contohnya seperti role play game (rpg), sport, simulasi, card game, shooting dan masih banyak lagi. Industri game pun telah banyak berkembang selama beberapa tahun terakhir.

Maka dari itu, peneliti tertarik untuk mengangkat topik ini ke dalam sebuah penelitian. Yang ingin dicari tahu adalah genre yang paling banyak dimainkan, serta rata-rata dari jam bermain bagi orang-orang yang bermain game. Penelitian ini diharapkan nantinya dapat menjadi patokan untuk developer game untuk membuat game yang menarik minat dari gamers berdasarkan genre dan kategori game.

B. TUJUAN PROJEK

Kita sebagai company game ingin memulai project game baru dan ingin memulai dengan ide pembuatan game. Dengan riset kepopuleran genre game dapat membantu kita untuk membuat game bergenre apa atau game seperti apa.

1. H_0 : Tidak ada korelasi dalam preferensi genre game dengan jenis kelamin pemilih

H_a : Terdapat korelasi dalam preferensi genre game dengan jenis kelamin pemilih

2. H_0 : Rata-rata data lama bermain sama dengan nilai tengah confidence interval

H_a : Rata-rata data lama bermain lebih rendah atau lebih tinggi dari nilai tengah confidence interval

3. H_0 : Tidak ada hubungan antara durasi bermain (X) dan durasi istirahat (Y)

H_a : Terdapat hubungan positif antara durasi bermain (X) dan durasi istirahat (Y)

II. KAJIAN LITERATUR

Dalam arti yang lebih luas, istilah "permainan" juga dapat merujuk pada aktivitas apa pun yang menyenangkan atau menantang, dan melibatkan seperangkat aturan dan tujuan. Ini dapat mencakup kegiatan seperti teka-teki, tantangan, dan kompetisi. Secara keseluruhan, game adalah aktivitas yang

populer dan dinikmati secara luas yang dapat digunakan untuk hiburan, pembelajaran, dan pengembangan pribadi.

Pada analisis digunakan metode statistika deskriptif dan statistika inferensial. Statistika deskriptif adalah cabang dari statistika yang bertujuan untuk menggambarkan, menjelaskan, dan memahami data yang ada. Statistika deskriptif menggunakan teknik-teknik kuantitatif untuk mengumpulkan, mengelompokkan, menganalisis, dan menyajikan data dalam bentuk yang bisa dipahami oleh orang lain. Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu data sehingga memberikan informasi yang berguna (Walpole, 1995).

Metode statistika deskriptif yang digunakan pada analisis ini adalah metode pemusatan data yang mana berisi berupa:

1. Tendensi Central: mean, modus, dan median.

2. Posisi/letak data : kuartil

3. Penyebaran data : standar deviasi varians

4. Asosiasi data : kovarians, korelasi

5. Bentuk data : skewness, kurtosis

Mean adalah nilai rata-rata dari sekumpulan data. Mean dihitung dengan menjumlahkan semua nilai dalam sekumpulan data, kemudian membagi jumlah tersebut dengan jumlah data yang ada. rumus mean :

$$\bar{x} = \frac{\sum x_i}{n}$$

Keterangan :

\bar{x} = Nilai rata-rata

x_i = Nilai sample ke - i

n = Jumlah Sampel

B_m = Bonderi bawah dari kelas median
 $(\sum f_i)_0$ = Jumlah frekuensi dari kelas -

kelas sebelumnya
 f_m = frekuensi dari kelas median

I = interval dari kelas median

Kuartil adalah nilai yang membagi sekumpulan data menjadi empat bagian

yang sama besar (atau setengah dari median). Ada tiga kuartil yang dapat ditentukan dari sekumpulan data, yaitu:

1. Kuartil pertama (Q1) adalah nilai yang membagi sekumpulan data menjadi dua bagian yang sama besar, dengan bagian yang lebih kecil terdiri dari 25% data terkecil.
2. Kuartil kedua (Q2) adalah nilai median dari sekumpulan data.
3. Kuartil ketiga (Q3) adalah nilai yang membagi sekumpulan data menjadi dua bagian yang sama besar, dengan bagian yang lebih besar terdiri dari 25% data terbesar.

rumus kuartil yaitu :

$$Q_1 = x_{\frac{n+1}{4}}$$

$$Q_2 = x_{\frac{2(n+1)}{4}}$$

$$Q_3 = x_{\frac{3(n+1)}{4}}$$

Varians adalah ukuran spread atau dispersi dari sekelompok data. Ini mengukur seberapa jauh setiap nilai dari

rata-rata (mean) dari sekelompok data.

rumus varians yaitu :

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - Me)^2}{N - 1}$$

σ = variansi

i = bilangan asli

x_i = data tengah untuk setiap kelas

Me = mean (nilai rata - rata)

N = banyak data

Standard deviation adalah ukuran variabilitas suatu distribusi data. Standar Deviasi menunjukkan seberapa jauh rata-rata nilai-nilai dari distribusi tersebut terpisah dari nilai rata-ratanya. Rumus standar deviasi yaitu:

Data tunggal

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

S = Standar Deviasi

x_i = nilai x ke-i

\bar{x} = nilai rata - rata data

n = jumlah data

Data kelompok

$$S = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{n}}$$

S = Standar Deviasi

f_i = frekuensi Kelompok

i

x_i = nilai x ke-i

\bar{x}

= nilai rata - rata data

Kovariansi adalah ukuran korelasi antara dua variabel. Nilai kovariansi positif menunjukkan bahwa kedua variabel tersebut bergerak ke arah yang sama. rumus Kovariansi yaitu :

$$\Sigma (x - \bar{x})(y - \bar{y})$$

$$cov_{x,y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N - 1}$$

$cov_{x,y}$ = kovarian antara variabel x dan y

x_i = data value of x

y_i = data value of y

\bar{x} = mean of x

\bar{y} = mean of y

N = number of data values

Koefisien korelasi merupakan

ukuran kuantitatif dari tingkat korelasi untuk beberapa tipe korelasi (hubungan), yang menyatakan hubungan secara statistik antara dua variabel. Nilai koefisien korelasi pearson bernilai antara -1 sampai dengan 1. rumus korelasi pearson yaitu :

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

r = koefisien korelasi pearson

x_i = data x

y_i = data y

Skewness adalah ukuran statistik

simetri dari sebuah distribusi. Ini mewakili

terdistorsi ke kanan jika memiliki ekor panjang di sisi kanan dan ekor yang lebih pendek di sisi kiri. Skewness penting dalam analisis statistik karena dapat mempengaruhi validitas uji statistik dan interpretasi hasilnya

rumus skewness yaitu :

$$SK = \frac{\mu - Mo}{\sigma}$$

$$SK = \frac{3(\mu - Md)}{\sigma}$$

SK = Koefisien Skewness

μ = nilai rata - rata hitung

Mo = nilai modus

σ = Standar Deviasi

Statistika Inferensial Merupakan semua metode yang berhubungan dengan analisis sebagian data untuk peramalan atau penarikan kesimpulan mengenai keseluruhan gugus data induknya. (Walpole,1995). Metode statistika

inferensial yang digunakan dalam penelitian ini adalah uji hipotesis. Uji hipotesis adalah proses yang digunakan untuk menguji kebenaran atau kevalidan suatu hipotesis yang telah diajukan. Hipotesis adalah sebuah asumsi atau dugaan tentang sesuatu yang belum diketahui dengan pasti. Uji hipotesis yang digunakan adalah menguji confidence interval pada satu populasi dan dua sejauh mana distribusi tersebut terdistorsi atau tidak simetris.

Rumus
Confidence
Interval :

$$CI = \bar{x} \pm z \frac{s}{\sqrt{n}}$$

CI = confidence interval

\bar{x} = sample mean

z = confidence level value

s = sample standard deviation

III. METODOLOGI DAN IMPLEMENTASI

A. Objek Penelitian

Objek yang dipilih adalah genre yang dimainkan dari data yang diambil dari google form. Alasan dipilihnya objek ini adalah untuk mengetahui game paling favorit yang dimainkan mahasiswa umn sebagai patokan untuk game developer dalam membuat game. Dengan membuat pertanyaan tentang umur, dan kategori yang dimainkan, peneliti dapat mengetahui patokan umur dari responden.

B. Metode Penelitian

Penelitian dilakukan dengan aplikasi Rstudio dan aplikasi microsoft excel. Setelah menyebarkan google form, dilakukan pengolahan dan pengelompokan

data melalui microsoft excel untuk memudahkan pemrosesan data.

Selanjutnya, Peneliti Menganalisis

dan merekap data yang telah dirapikan melalui aplikasi Rstudio supaya dapat diproses dengan mudah. Total data yang didapatkan adalah 89 data yang diisi oleh responden dari mahasiswa universitas multimedia nusantara dengan rentan umur 18-22 tahun.

C. Data Penelitian

Data penelitian didapatkan dari google form yang memuat pertanyaan tentang umur, angkatan dan genre favorit responden. Dari data yang diambil, data yang peneliti kelola adalah data tentang genre yang berupa data numerik. Dari data numerik tersebut didapatkan data sebagai berikut:

- (1) Rata-Rata waktu bermain game Mahasiswa Universitas Multimedia Nusantara adalah 3 jam lebih.
- (2) Game dengan Genre Action merupakan genre yang paling banyak dimainkan oleh Mahasiswa Universitas Multimedia Nusantara. Rata-Rata waktu Durasi Rehat dari bermain game adalah 8 jam lebih.

D. Metode Pengumpulan Data

Peneliti menggunakan data primer sebagai sumber data, yang diambil dari melakukan survei melalui google form

kepada mahasiswa universitas multimedia nusantara. Data primer adalah data pertama kali yang dikumpulkan oleh peneliti melalui upaya pengambilan data di lapangan langsung. Dari data yang telah diperoleh peneliti mengambil data umur dan genre yang disukai mahasiswa universitas multimedia nusantara. Tipe data tersebut adalah data dependen artinya data yang mempengaruhi atau dipengaruhi oleh data lain.

E. Metode Analisis Data

Metode yang dipakai dalam penelitian ini adalah statistika deskriptif dan statistika inferensi. Dalam statistika deskriptif yang dipakai adalah :

1. Tendensi Sentral: mean, modus, dan median.
 2. Posisi/letak data : kuartil
 3. Penyebaran data : standar deviasi, varians
 4. Asosiasi data : kovarians, korelasi
 5. Bentuk data : skewness, kurtosis
- tujuan penggunaan statistika inferensi adalah agar data lebih mudah diolah.

Lalu dalam statistika inferensi yang dipakai adalah uji hipotesis dimana uji hipotesis yang dilakukan adalah dengan menguji confidence interval pada satu populasi dan dua populasi.[15].

IV. HASIL DAN ANALISIS

A. Mean, Median, Modus, kurtosis, standard deviasi

Berikut akan ditampilkan beberapa gambar untuk memberikan informasi terkait hasil dari analisis yang telah dilakukan. Untuk analisis pertama akan menjelaskan tentang hasil dari central tendencies yang berisikan *mean*, *modus*, *median*, *kurtosis* serta *standard deviation* dari data yang berhasil dikumpulkan.

Berdasarkan data yang telah diberikan, terdapat beberapa kesimpulan yang dapat diambil mengenai tendensi central dari data lama_durasibermain:

```
> mean(lama_durasibermain)
[1] 3.876404
> median(lama_durasibermain)
[1] 3
>
> modus <- function(x){
+   uniqx <- unique(x)
+   uniqx[which.max(tabulate(match(x, uniqx)))]
+ }
> modus(lama_durasibermain)
[1] 3
>
> #variability
> range(lama_durasibermain)
[1] 1 10
> var(lama_durasibermain)
[1] 4.973187
> sd(lama_durasibermain)
[1] 2.230064
>
> #Untuk menghitung relative standing
> quantile(lama_durasibermain)
 0%  25%  50%  75% 100%
  1    3    3    4   10
>
> #Untuk menentukan bentuk (shape)
> summary(lama_durasibermain)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 1.000  3.000   3.000   3.876  4.000 10.000
> skewness(lama_durasibermain)
[1] 1.207087
> kurtosis(lama_durasibermain)
[1] 3.922841
```

Mean (rata-rata) dari data lama_durasibermain adalah 3.87. Mean adalah nilai rata-rata dari seluruh data, dihitung dengan menjumlahkan seluruh nilai data dan membagi dengan jumlah data. Median dari data lama_durasibermain adalah 3. Median

adalah nilai tengah dari sekumpulan data yang telah diurutkan secara ascending.

Modus yang telah di dapat kan dari data lama_durasibermain adalah 3. Modus adalah nilai yang paling sering muncul dalam sekumpulan data.

Range dari data lama_durasibermain adalah 1-10. Range adalah selisih antara nilai tertinggi dan terendah dalam sekumpulan data.

Variance dari data lama_durasibermain adalah 4.97. Variance adalah ukuran variasi dari sekumpulan data, yang dihitung dengan menghitung selisih antara setiap data dengan mean, kemudian dikuadratkan dan dibagi dengan jumlah data. Standar deviasi dari data lama_durasibermain adalah 2.23. Standar deviasi adalah ukuran variasi dari sekumpulan data, yang dihitung dengan menghitung akar kuadrat dari variance. Quantile dari data lama_durasibermain menunjukkan bahwa 25% data memiliki nilai di bawah 3, 50% data memiliki nilai di bawah 3, dan 75% data memiliki nilai di bawah 4. Quantile adalah nilai di bawah mana sejumlah persentase tertentu dari sekumpulan data terletak.

Skewness dari data lama_durasibermain adalah 1.21.

Kurtosis dari data lama_durasibermain adalah 3.92. Kurtosis adalah ukuran kelebihan atau kekurangan puncak distribusi data dibandingkan dengan distribusi normal.

Selanjutnya adalah Analisis dan kesimpulan Bentuk Statistic Deskriptif dari lama_durasiistirahat

```
> #ukuran Central Tendency
> mean(lama_durasiistirahat)
[1] 5.202247
> median(lama_durasiistirahat)
[1] 5
> modus <- function(x){
+   uniqx <- unique(x)
+   uniqx[which.max(tabulate(match(x, uniqx)))]
+ }
> modus(lama_durasiistirahat)
[1] 7
> #variability
> range(lama_durasiistirahat)
[1] 1 10
> var(lama_durasiistirahat)
[1] 9.276813
> sd(lama_durasiistirahat)
[1] 3.045786
>
> #Untuk menghitung relative standing
> quantile(lama_durasiistirahat)
 0% 25% 50% 75% 100%
 1   2   5   7  10
> #Untuk menentukan bentuk (shape)
> summary(lama_durasiistirahat)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
1.000   2.000   5.000   5.202   7.000  10.000
> skewness(lama_durasiistirahat)
[1] 0.04446785
> kurtosis(lama_durasiistirahat)
[1] 1.708646
> summary(lama_durasiistirahat)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
1.000   2.000   5.000   5.202   7.000  10.000
```

Dari data lama_durasiistirahat yang disajikan, terlihat bahwa nilai rata-rata (mean) adalah 5.202247, nilai tengah (median) adalah 5, dan modus adalah 7. Ini menunjukkan bahwa data memiliki cenderung ke arah nilai tengah dan nilai modus. Nilai tertinggi yang tercatat dalam data adalah 10, sedangkan nilai terendah adalah 1, sehingga rentang data adalah

$10 - 1 = 9$. Nilai variansi (var) adalah 9.276813 dan nilai standar deviasi (sd) adalah 3.045786, yang menunjukkan bahwa data tersebar cukup luas di sekitar nilai rata-rata.

Jika kita melihat nilai quantile, terlihat bahwa 25% data berada di bawah nilai 2, 50% data berada di bawah nilai 5, dan 75% data berada di bawah nilai 7. Ini menunjukkan bahwa distribusi data cenderung menyebar di sekitar nilai tengah.

Dari hasil ringkasan (summary) data, terlihat bahwa nilai minimum adalah 1, nilai terkecil yang kelima (1st quartile) adalah 2, nilai tengah (median) adalah 5, nilai rata-rata (mean) adalah 5.202, nilai terbesar yang keempat (3rd quartile) adalah 7, dan nilai maksimum adalah 10. Ini menunjukkan bahwa data memiliki cenderung ke arah nilai tengah, namun memiliki beberapa nilai yang lebih tinggi atau lebih rendah daripada nilai tengah.

Nilai skewness adalah 0.04446785, yang menunjukkan bahwa data tidak terlalu terdistorsi ke satu sisi atau yang lain. Nilai kurtosis adalah 1.708646, Kesimpulannya, data lama_durasiistirahat memiliki cenderung ke arah nilai tengah dan tidak terlalu terdistorsi ke sisi kiri atau kanan.

shapiro-wilk normality test

```
data: lama_durasibermain
W = 0.84585, p-value = 3.456e-08
> 3,717 4,03584
```

data lama durasi bermain di uji dengan uji shapiro-wilk mendapatkan p value yang terbilang sangat kecil dimana p-value berjumlah $3,456 \times 10^{-8}$. Hal p value terbukti lebih kecil dari α yang menandakan bahwa penyebaran data tersebar dengan tidak normal. hal ini juga ditandakan dengan banyaknya data yang tidak mengenai garis lurus pada gambar di bawah.

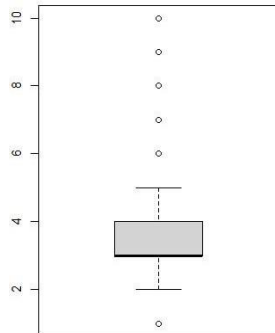
B. Korelasi dan Kovariansi

```
[1] 0.0034023
> cor(lama_durasibermain,lama_durasiistirahat)
[1] 0.8854023
> cov(lama_durasiistirahat,lama_durasibermain)
[1] 6.013917
```

Dari hasil analisis yang Anda berikan, terdapat korelasi positif yang kuat antara lama durasi bermain dengan lama durasi istirahat. Hal ini dapat dilihat dari nilai korelasi Pearson yang bernilai 0.8854023, yang menunjukkan bahwa semakin lama durasi bermain, maka semakin lama pula durasi istirahat yang diambil. Nilai kovariansi yang bernilai 6.013917 juga menunjukkan adanya hubungan positif antara kedua variabel tersebut.

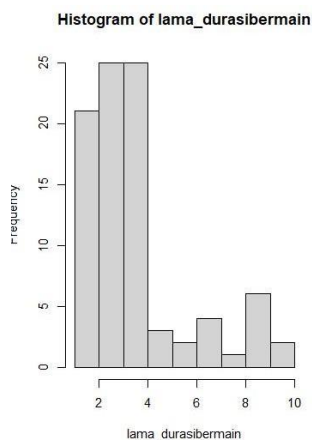
Kesimpulan yang dapat diambil dari data tersebut adalah bahwa semakin lamaseseorang bermain, maka semakin lama pula ia akan mengambil istirahat.

C. Grafik

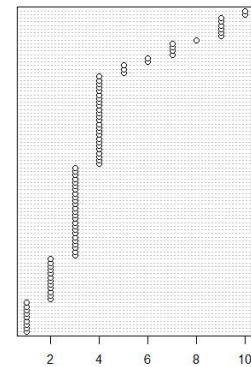


Dari grafik boxplot Lama_durasibermain didapatkan nilai minimum angka 2 dengan nilai maksimum angka 5, Lalu kuartil pertama didapatkan angka 3 dengan nilai kuartil tiga didapatkan angka 4. Median yang didapatkan adalah angka 3

Berikut merupakan grafik histogram dari lama_durasibermain yang didapatkan data berupa :



Berikut merupakan grafik Dot dari lama_durasibermain yang didapatkan data berupa :



Berikut merupakan syntax dari grafik boxplot, histogram, dan grafik dot dari lama_durasibermain

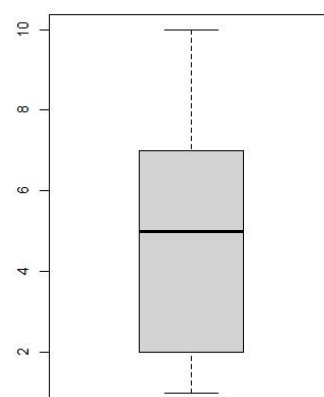
```
#nomor3
#Instalasi Package
install.packages("ggplot2")

#lama_durasibermain
#Diagram Boxplot
boxplot(lama_durasibermain)

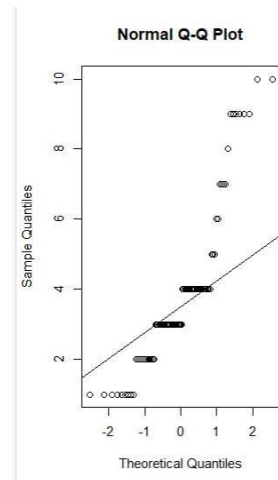
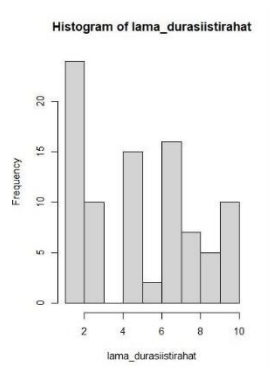
#Diagram Histogram
hist(lama_durasibermain)

#Diagram dotchart
dotchart(lama_durasibermain)
```

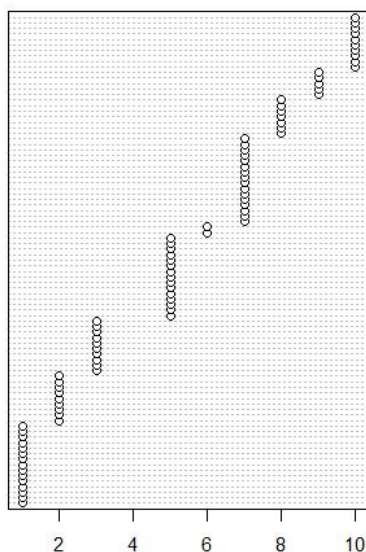
Di Dalam syntax tersebut, dimasukan dulu data yang sudah diambil dari google form



Berikut merupakan grafik histogram dari lama_durasiistirahat yang didapatkan data berupa :

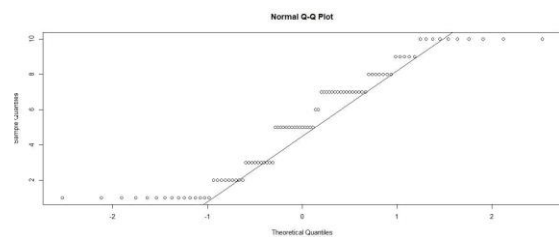


Berikut merupakan grafik dot dari lama_durasiistirahat yang didapatkan data berupa :



```
> #LAMA DURASI BERMAIN  
> lama_durasi_bermain <- c(1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,  
3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,4,4,4,4,4,4,4,4,4,4,4,4,  
4,4,4,4,4,4,4,4,5,5,5,6,6,7,7,7,8,9,9,9,9,9,9,10,10)  
> qqnorm(lama_durasi_bermain)  
> qqline(lama_durasi_bermain)
```

Grafik plot Lama_durasibermain



```
> lama_durasiistirahat <-c(1,1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,3,
3,3,3,3,3,3,3,5,5,5,5,5,5,5,5,5,5,5,5,6,6,7,7,7,7,7,7,7,7,7,7,
7,7,7,7,8,8,8,8,8,8,9,9,9,9,9,10,10,10,10,10,10,10,10)
> qqnorm(lama_durasiistirahat)
> qqline(lama_durasiistirahat)
```

Grafik plot Lama_durasibelajar

Plot QQ (Quantile-Quantile plots) adalah alat grafis yang digunakan untuk menilai apakah sekumpulan data mengikuti distribusi tertentu. Mereka berguna untuk membandingkan distribusi kumpulan data

D. Grafik Normal Q-Q Plot

dengan distribusi teoretis, seperti distribusi normal.

Dalam hal ini, plot QQ digunakan untuk membandingkan distribusi data lama_durasibermain dengan distribusi normal. Plot QQ menampilkan sebaran data, dengan garis yang ditarik melewatinya yang mewakili distribusi normal. Jika data mengikuti distribusi normal, titik-titik pada scatterplot akan mendekati garis.

Untuk menginterpretasikan plot QQ, Anda harus melihat pola titik-titik pada plot tersebut. Jika titik-titik tersebut membentuk garis lurus, hal ini menunjukkan bahwa data mengikuti distribusi normal. Jika titik-titik tersebut menyimpang jauh dari garis, hal ini menunjukkan bahwa data tidak mengikuti distribusi normal.

Dalam hal ini, plot QQ menunjukkan bahwa data tidak mengikuti distribusi normal. Titik-titik pada scatterplot membentuk kurva, bukan garis lurus, yang menunjukkan bahwa data tidak terdistribusi secara normal. Hal ini menunjukkan bahwa sebaran data lama_durasibermain berbeda dengan sebaran normal.

Sebagai kesimpulan, plot QQ menunjukkan bahwa data lama_durasibermain tidak mengikuti distribusi normal. Hal ini menunjukkan bahwa sebaran data berbeda dengan sebaran normal.

E. Confidence Interval

```
#input sample size, sample mean, and sample standard deviation
n <- 8
xbar <- 28.25
s <- 10.61

#calculate margin of error
margin <- qt(0.50,df=n-1)*s/sqrt(n)

#calculate lower and upper bounds of confidence interval
low <- xbar - margin
low
[1] 25.72

high <- xbar + margin
high
[1] 30.78
```

Berdasarkan uji sampel data genre yang dilakukan dengan jumlah sampel sebanyak 8, diperoleh nilai rata-rata sampel sebesar 28.25 dan standar deviasi sampel sebesar 10.61. Dengan menggunakan tingkat kepercayaan 50%, interval kepercayaan yang diperoleh adalah (25.72, 30.78). Ini berarti bahwa dengan tingkat kepercayaan 50%, nilai rata-rata populasi terletak di antara batas atas dan batas bawah interval kepercayaan tersebut. Oleh karena itu, kita tidak dapat dengan yakin menyatakan nilai rata-rata populasi sebenarnya, tetapi dapat memberikan batasan untuk nilai tersebut berdasarkan sampel yang tersedia.

Secara lebih spesifik, interval kepercayaan merupakan estimasi yang memperkirakan nilai rata-rata populasi dengan tingkat kepercayaan tertentu (dalam kasus ini adalah 50%). Semakin kecil jumlah sampel yang

digunakan, semakin besar margin of error yang dihasilkan, sehingga semakin lebar interval kepercayaan yang diperoleh. Sebaliknya, semakin besar jumlah sampel yang digunakan, semakin kecil margin of error yang dihasilkan, sehingga semakin sempit interval kepercayaan yang diperoleh.

```
#NOMOR 6
#2 sample
#lama bermain dan lama durasi istirahat
set.seed(123) # Setting seed for reproducibility
lama_bermain <- rnorm(89, mean = 3.876404, sd = 2.2)
lama_durasiistirahat <- rnorm(89, mean = 5.202247, sd = 2.450324)

# Perform an independent two-sample t-test
t_test_result <- t.test(lama_bermain, lama_durasiistirahat)

# Display the results
cat("Two-Sample t-test:\n")
print(t_test_result)

# Confidence interval for the difference between means
conf_interval <- t_test_result$conf.int
cat("\n95% Confidence Interval for the Difference Between Means:")
cat("\nLower Bound:", conf_interval[1])
cat("\nUpper Bound:", conf_interval[2])

welch Two Sample t-test

data: lama_bermain and lama_durasiistirahat
t = -3.4842, df = 166.79, p-value = 0.0006305
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.8311229 -0.5065315
sample estimates:
mean of x mean of y
 3.963436  5.132263

> # Confidence interval for the difference between means
> conf_interval <- t_test_result$conf.int
> cat("\n95% Confidence Interval for the Difference Between Means:")
95% Confidence Interval for the Difference Between Means:
> cat("\nLower Bound:", conf_interval[1])
Lower Bound: -1.831123
> cat("\nUpper Bound:", conf_interval[2])
Upper Bound: -0.5065315
`
```

Hasil dari uji t-test dua sampel dan interval kepercayaan memberikan informasi tentang perbandingan antara dua populasi (lama_bermain dan lama_durasiistirahat). Mari kita interpretasikan hasilnya:

Uji t-Test Dua Sampel:

Nilai t: -3.4842

Derajat Kebebasan (df): 166.79

Nilai p: 0.0006305

Uji t digunakan untuk menilai apakah terdapat perbedaan yang signifikan antara rata-rata dua populasi. Nilai t adalah -

3.4842, dan derajat kebebasan sekitar 167. Nilai p sangat kecil (0.0006305), yang kurang dari tingkat signifikansi tipikal 0.05. Oleh karena itu, kita menolak hipotesis nol.

Hipotesis nol (H_0) dalam hal ini adalah bahwa tidak ada perbedaan antara rata-rata lama_bermain dan lama_durasiistirahat. Nilai p kecil menunjukkan bahwa ada cukup bukti untuk menyimpulkan bahwa perbedaan yang signifikan ada.

Interval Kepercayaan 95% untuk Perbedaan Rata-rata:

Batas Bawah: -1.8311229

Batas Atas: -0.5065315

Interval kepercayaan memberikan rentang nilai yang mungkin untuk perbedaan sejati antara rata-rata. Dalam hal ini, kita 95% yakin bahwa perbedaan sejati antara rata-rata lama_bermain dan lama_durasiistirahat berada di antara -1.8311229 dan -0.5065315.

Secara ringkas, berdasarkan hasil uji t-test dan interval kepercayaan, terdapat bukti kuat yang menunjukkan bahwa rata-rata lama_bermain secara signifikan berbeda dari rata-rata lama_durasiistirahat.

F. Linear regresi

Dari hasil analisis data yang telah Anda berikan, dapat dilihat bahwa terdapat hubungan linier yang kuat antara lama durasi istirahat (y) dengan lama durasi bermain (x). Hal ini dapat dilihat dari nilai koefisien x dari model regresi linier yang bernilai 1.2093 yang signifikan secara statistik ($p\text{-value} < 2.2e-16$). Ini menunjukkan bahwa setiap kenaikan satu satuan pada lama durasi bermain (x), akan menyebabkan kenaikan sebesar 1.2093 satuan pada lama durasi istirahat (y).

Nilai R-squared yang bernilai 0.7839 juga menunjukkan bahwa model regresi linier ini mampu menjelaskan sekitar 78% variasi dari lama durasi istirahat (y) yang terjadi. Sisanya sekitar 22% dijelaskan oleh faktor-faktor lain yang tidak diteliti dalam model ini.

Untuk memprediksi lama durasi istirahat (y) untuk lama durasi bermain (x) tertentu, Anda dapat menggunakan fungsi `predict` dengan menentukan nilai-nilai x yang ingin diprediksi. Hasilnya adalah lama durasi istirahat (y) yang diperkirakan oleh model regresi linier tersebut.

Residu adalah selisih antara nilai y yang sebenarnya dengan nilai y yang diperkirakan oleh model regresi linier. Jika selisih ini semakin kecil, berarti model regresi linier tersebut semakin

baik dalam memprediksi nilai y . Dari hasil residu yang Anda berikan, terlihat bahwa

[illegible]

selisih antara nilai y yang sebenarnya dengan nilai y yang diperkirakan oleh model regresi linier tersebut cukup kecil, yang menunjukkan bahwa model regresi linier tersebut cukup baik dalam memprediksi lama durasi istirahat (y).

Secara keseluruhan, dapat disimpulkan bahwa terdapat hubungan linier yang kuat antara lama durasi bermain (x) dengan lama durasi istirahat (y), dan model regresi linier yang dibangun cukup baik dalam memprediksi lama durasi istirahat (y) untuk lama durasi bermain (x) tertentu.

G. Pengujian Chi square terhadap data Kategorik

```
> print(contingency_table)
```

	Pria	wanita
Action	46	0
Fighting	24	0
Sport	27	0
Racing	21	0
Shooting	32	6
Strategy	16	16
Role Playing Game RPG	23	16
Adventure	13	12
Simulasi	10	10
Card Game	7	6

```
> chi_square_result <- chisq.test(contingency_table)
```

Berdasarkan data diatas terlihat bahwa ada 10 jenis Game dan kita akan menyimpulkan, apakah terdapat korelasi antara genre game dan jenis kelamin atau tidak.

H(0) = Tidak ada korelasi antara genre game dan jenis kelamin pemilih.

H(1) = Terdapat korelasi antara genre game

dan jenis kelamin pemilih.

```
> print(chi_square_result)
Pearson's chi-squared test
data: contingency_table
X-squared = 77.305, df = 9, p-value = 5.532e-13
> if (chi_square_result$p.value < 0.05) {
+   cat("Terdapat bukti statistik yang cukup untuk menolak hipotesis nol.\n")
+ } else {
+   cat("Tidak terdapat bukti statistik yang cukup untuk menolak hipotesis nol.\n")
+ }
Terdapat bukti statistik yang cukup untuk menolak hipotesis nol.
>
```

Kemudian, dilakukan uji chi-square untuk mengetahui apakah ada korelasi antara genre game dan jenis kelamin pemilih. Berdasarkan hasil uji chi square diatas, kita mendapatkan nilai ($p\text{-value} = < 0,5$), maka $H(0)$ ditolak dan kita mendapatkan kesimpulan bahwa terdapat korelasi yang signifikan antara genre game dan jenis kelamin pemilih.

V.PENUTUP

Dari hasil penelitian yang dilakukan, didapatkan hasil sebagai berikut:

1. Uji hipotesis utama menunjukan bahwa terdapat korelasi yang kuat antara genre game dengan jenis kelamin pemilih, dimana pemilih dengan jenis kelamin pemilih pria memiliki preferensi genre game yang berbeda dengan pemilih jenis kelamin wanita, dimana $H(0)$ ditolak karena mendapatkan ($p\text{-value} = < 0,5$) dari hasil uji chi square
2. Uji hipotesis menunjukan semakin lama durasi bermain, maka semakin lama pula durasi istirahat yang diambil. Nilai

kovariansi yang bernilai 6.013917 juga menunjukkan adanya hubungan positif antara kedua variabel tersebut.

3. Uji hipotesis menunjukkan bahwa rata - rata data lama bermain tidak melebihi dan kurang dari hasil confidence interval dimana, rata - rata lama bermain 3.8764 dan confidence interval (3.717, 4.03584).

DAFTAR PUSTAKA

- [1] Walpole, Ronald E. 1993. Pengantar Statistika Edisi Ke-3. Jakarta: PT Gramedia Pustaka Utama.
- [2] Carolina, G. (2021). Pengantar Konsep Statistika Deskriptif. zenius.net. <https://www.zenius.net/blog/statistika-deskriptif>.
- [3] Wibowo, A. (2017). Uji Chi-Square pada statistika dan SPSS. *Jurnal Ilmiah SINUS*, 4(2).
- [4] Nasrum, A. (2018). Uji normalitas data untuk penelitian. *Jayapangus Press Books*, i-117.
- [5] Quraisy, A. (2020). Normalitas Data Menggunakan Uji Kolmogorov-Smirnov dan Saphiro-Wilk: Studi kasus penghasilan orang tua mahasiswa Prodi Pendidikan Matematika Unismuh Makassar. *Journal of Health Education Economics Science and Technology (J-HEST)*, 3(1), 7-11.
- [6] Kronthaler, F., & Zöllner, S. (2021). Data analysis with RStudio. *An Easygoing Introduction*, 7-131.
- [7] Artusi, R., Verderio, P., & Marubini, E. (2002). Bravais-Pearson and Spearman correlation coefficients: meaning, test of hypothesis and confidence interval. *The International journal of biological markers*, 17(2), 148-151.
- [8] Horton, N. J., & Kleinman, K. (2015). *Using R and RStudio for data management, statistical analysis, and graphics*. CRC Press.
- [9] Ramachandran, K. M., & Tsokos, C. P. (2020). *Mathematical statistics with applications in R*. Academic Press.
- [10] Van der Loo, M. P. (2012). *Learning RStudio for R statistical computing*. Packt Publishing Ltd.
- [11] Asdi, Y. (2017). Pengenalan Software R. Website: https://nanopdf.com/download/pengenalan-software-r_pdf, diakses, 7.
- [12] Heiberger, R. M., Heiberger, R. M., & Burt Holland, B. H. (2015). *Statistical Analysis and Data Display An Intermediate Course with Examples in R*. Springer.
- [13] Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in ecology and evolution*, 4(8), 782-788.
- [14] Schumacker, R., & Tomek, S. (2013). *Understanding statistics using R*. Springer Science & Business Media.
- [15] Vivi Silvia, S. E. (2020). *Statistika Deskriptif*. Penerbit Andi.