**LAPORAN PENGERJAAN TUGAS II**

**MATA KULIAH PROBABILITAS DAN STATISTIKA**



Disusun Oleh:

Alfi Nor Ihsan

2109106018

**PROGRAM STUDI INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS MULAWARMAN**

**SAMARINDA**

**2022**

**KATA PENGANTAR**

Puji syukur kehadirat Allah SWT karena atas segala rahmat-Nya dan karunia-Nya lah, makalah ini dapat tersusun dengan baik dan tepat pada waktunya. Tidak lupa penulis mengucapkan terima kasih kepada bapak Dr. H. Fahrul Agus, S.Si, MT selaku dosen pengampu mata kuliah Probabilitas dan Statistika yang telah membimbing serta arahan, sehingga laporan ini dapat disusun dengan baik.

Adapun tujuan dari penyusunan laporan ini adalah untuk melaporkan pengerjaan tugas II di mata kuliah Probabilitas dan Statistika yang di tujukan untuk memenuhi nilai ulangan tengah semester pada mata kuliah Probabilitas dan Statistika.

Penulis menyadari bahwa laporan ini jauh dari kata sempurna dan masih terdapat memiliki banyak kekurangan dalam penyusunan laporan ini. Oleh karena itu penulis sangat mengharapkan saran dan kritik yang bersifat membangun dari pembaca untuk penyempurnaan makalah ini.

Samarinda, 06 Oktober 2022

Penyusun

**TUGAS II**

1. Buat Tabel Distribusi Frekuensi dan Histogram variabel reading\_score dan writing\_score dengan jumlah dan lebar interval sebesar 10.
2. Tentukan nilai Tendensi Central: Rata-rata, Median, Modus, Min dan Max data kedua variabel tersebut.
3. Tentukan nilai Variasi: Jangkauan Range, Kisaran Interkuartil, Varians, Standar Deviasi data *reading\_score* dan *writing\_score.*
4. Buat Grafik data Variabel reading\_score dan writing\_score: Histogram, BOX PLOT, dan Tentukan Nilai Skewness dan Kurtosis. Buat interpretasi saudara tentang perbedaan distribusi kedua variable tersebut, terutama pada gambar BOX PLOT dan nilai Skewness - Kurtosisnya.
5. Buat scater plot antara *writing\_score* (dependent variable-Y) versus *reading\_score* (independent variable-X). Tulis interpretasi saudara apakah ada indikasi hubungan linear antar keduanya.

**JAWABAN TUGAS II**

Pada pengerjaan tugas ini, saya menggunakan bahasa pemrograman Python untuk melakukan berbagai proses komputasi dan visualisasi, dan Visual Studio Code sebagai Text editor yang membantu saya dalam membuat dan menyelesaikan program ini. Adapun data yang pada tugas ini, sumber dataset yang di gunakan adalah “StudentsPerformance.csv”.

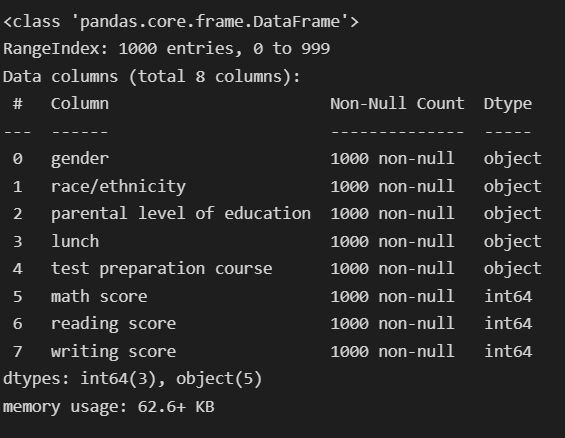
Adapun program yang akan saya lampirkan di bawah akan terpotong-potong agar penjelasan di tiap bagian program dapat di jelaskan, dan selain itu untuk mempermudah proses tersebut, saya menggunakan media Jupyter-lab pada penyelesaian kasus-kasus diatas.

Program tersebut di mulai dengan deklarasi library python yang di perlukan serta membaca data dari dataset.

|  |
| --- |
| # Deklarasi library python  import numpy as np  import pandas as pd  from prettytable import PrettyTable  from IPython.display import display\_html  import matplotlib.pyplot as plt |

|  |
| --- |
| # Membaca dataset  df = pd.read\_csv("StudentsPerformance.csv")  df.info() |

Adapun output dari proses pembacaan data diatas adalah:

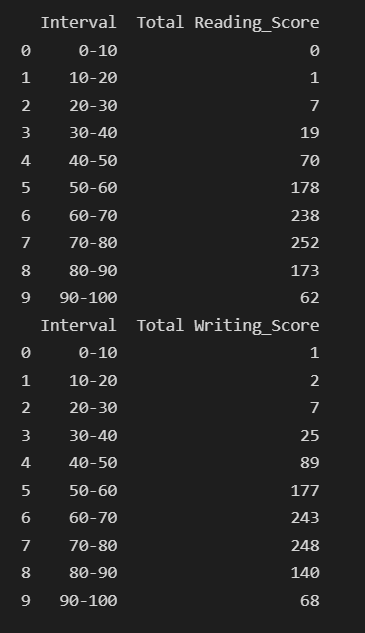


Selanjutnya merupakan penyelesaian dari tiap kasus sebelumnya yaitu :

1. **Buat Tabel Distribusi Frekuensi dan Histogram Variabel reading\_score dan writing\_score dengan jumlah dan lebar interval sebesar 10.**

Program yang di gunakan untuk membuat tabel distribusi dari variabel reading\_score dan reading\_score:

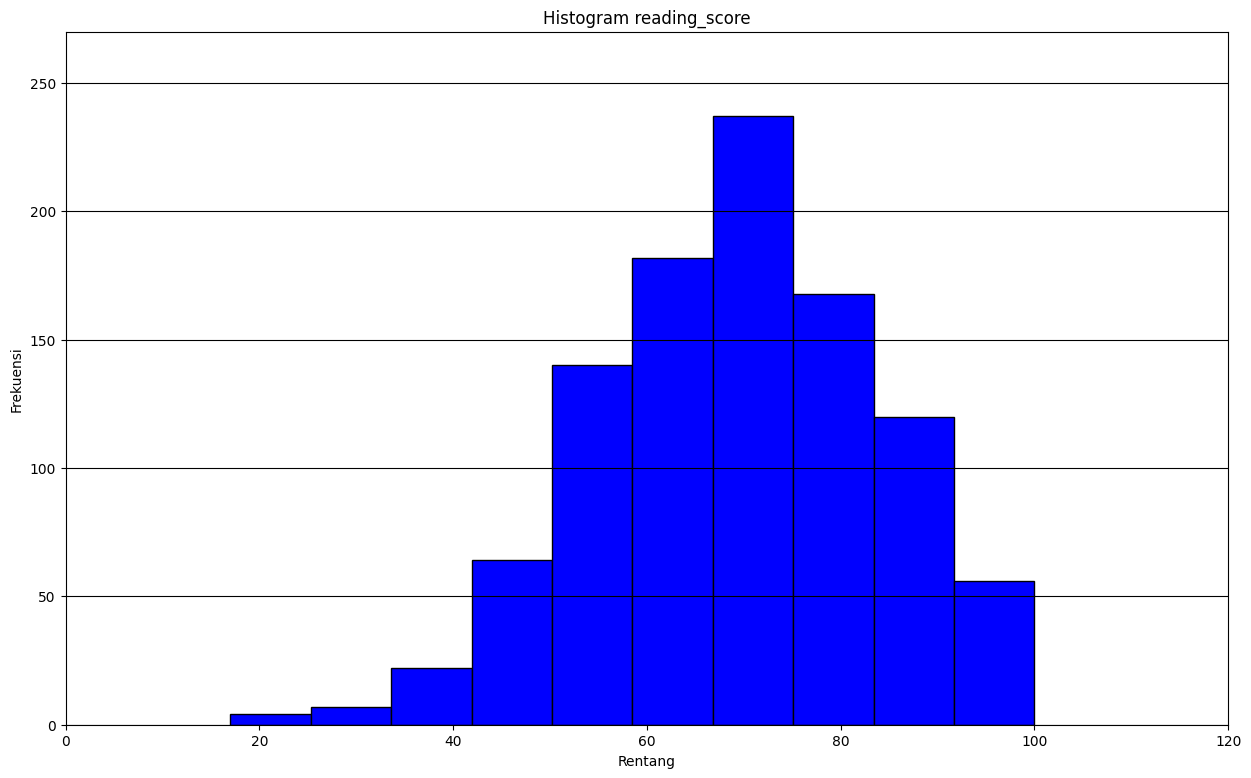
|  |
| --- |
| # Membuat tabel distribusi frekuensi dari Variabel reading\_score dan writing\_score.  #Interval dan keterangan  Interval = np.arange(0, 110, 10)  labels = np.vectorize(lambda x: f"{x}-{x + 10}")(np.arange(0, 100, 10))  #Nilai di tabel reading\_score  TotalReading\_score = df.groupby(pd.cut(df["reading score"], bins=Interval, labels=labels))["reading score"].count().to\_list()  TabelReading\_score = pd.DataFrame({ 'Interval': labels, 'Total Reading\_Score': TotalReading\_score })  #Nilai di setiap writing\_score  TotalWriting\_score = df.groupby(pd.cut(df["writing score"], bins=Interval, labels=labels))["writing score"].count().to\_list()  TabelWriting\_score = pd.DataFrame({ 'Interval': labels, 'Total Writing\_Score': TotalWriting\_score })  #menampilkan tabel reading\_score  print(TabelReading\_score)  #menampilkan tabel writing score  print(TabelWriting\_score) |

 Output dari program:

Program yang di gunakan untuk membuat histogram dari variabel reading\_score:

|  |
| --- |
| # Membuat histogram dari reading\_score  plt.figure(figsize=(15, 9))  plt.hist(  df["reading score"],  stacked=True,  color="blue",  edgecolor="black"  )  #ukuran dari rentang x dan y  plt.xlim(0, 120)  plt.ylim(0, 270)  plt.grid(axis='y', color='black')  # Keterangan pada histogram  plt.title("Histogram reading\_score")  plt.xlabel("Rentang")  plt.ylabel("Frekuensi")  # Menampilkan histogram reading\_score  plt.show() |

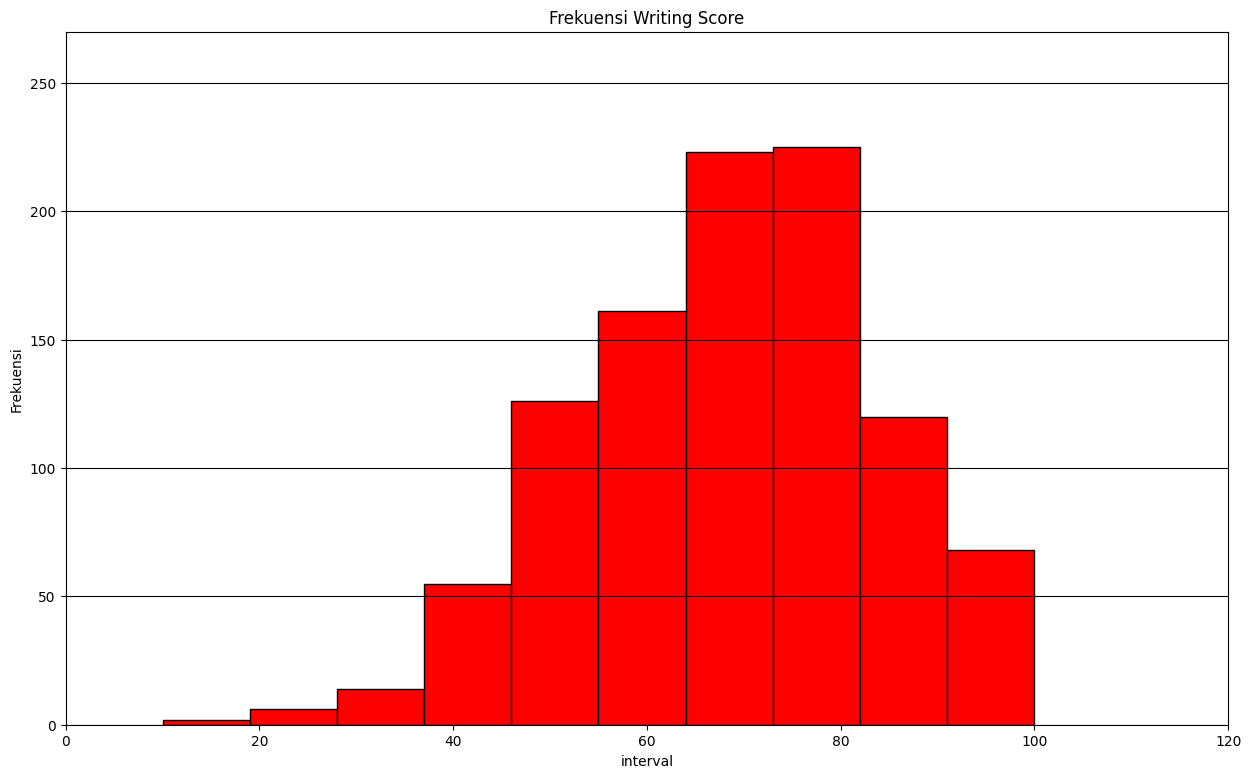
Dan output dari program tersebut adalah:



Program yang di gunakan untuk membuat histogram dari variabel writing\_score:

|  |
| --- |
| # Membuat histogram dari writing\_score  plt.figure(figsize=(15, 9))  plt.hist(  df["writing score"],  stacked=True,  color="red",  edgecolor="black"  )  #ukuran dari rentang x dan y  plt.xlim(0, 120)  plt.ylim(0, 270)  # Keterangan pada histogram  plt.title("Frekuensi Writing Score")  plt.xlabel("interval")  plt.ylabel("Frekuensi")  plt.grid(axis='y', color='black')  # Menampilkan histogram writing\_score  plt.show() |

Dan berikut output dari program diatas:

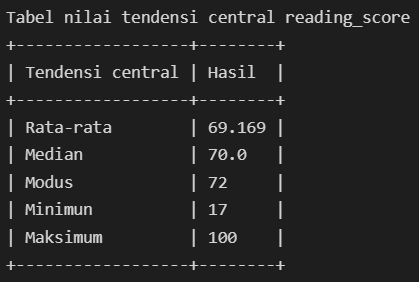


1. **Tentukan nilai Tendensi Central:** **Rata-rata, Median, Modus, Min dan Max data kedua variabel tersebut.**

Pada kasus ini, untuk menetukan nilai rata-rata, Median, Modus, Min dan Max data kedua variabel tersebut di perlukan sebuah tabel yang mana, di library Python sendiri sudah tersedia ***Prettytable*** untuk menampung data tersebut. Adapun Program yang di perlukan untuk menentukan Rata-rata, Median, Modus, Min dan Max data dari variabel reading\_score adalah:

|  |
| --- |
| # Mendapatkan nilai tendensi central (Mean, median, modus, min, max)  mean\_reading = df["reading score"].mean(axis = 0)  median\_reading = df["reading score"].median(axis = 0)  modus\_reading = df["reading score"].mode()[0]  min\_reading = df["reading score"].min(axis = 0)  max\_reading = df["reading score"].max(axis = 0)  # Membuat tabel tendensi central (Mean, median, modus, min, max)  tabel\_reading = PrettyTable(["Tendensi central", "Hasil"])  tabel\_reading.align = "l"  # Memuat nilai tendensi central (Mean, median, modus, min, max kedalam tabel  tabel\_reading.add\_row(["Rata-rata", mean\_reading])  tabel\_reading.add\_row(["Median", median\_reading])  tabel\_reading.add\_row(["Modus", modus\_reading])  tabel\_reading.add\_row(["Minimun", min\_reading])  tabel\_reading.add\_row(["Maksimum", max\_reading])  # Menampilkan tabel  print("Tabel nilai tendensi central reading\_score")  print(tabel\_reading) |

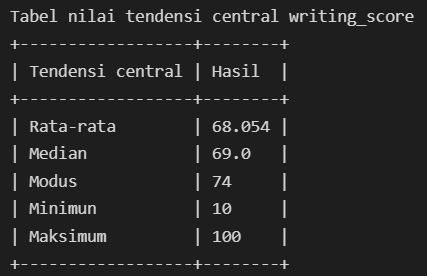
Dan Output yang di dapatkan dari program diatas adalah:



Program yang di perlukan untuk menentukan Rata-rata, Median, Modus, Min dan Max data dari variabel writing\_score adalah:

|  |
| --- |
| # Mendapatkan nilai tendensi central (Mean, median, modus, min, max)  mean\_writing = df["writing score"].mean(axis = 0)  median\_writing = df["writing score"].median(axis = 0)  modus\_writing = df["writing score"].mode()[0]  min\_writing = df["writing score"].min(axis = 0)  max\_writing = df["writing score"].max(axis = 0)  # Membuat tabel tendensi central (Mean, median, modus, min, max)  tabel\_writing = PrettyTable(["Tendensi central", "Hasil"])  tabel\_writing.align = "l"  # Memuat nilai tendensi central (Mean, median, modus, min, max kedalam tabel  tabel\_writing.add\_row(["Rata-rata", mean\_writing])  tabel\_writing.add\_row(["Median", median\_writing])  tabel\_writing.add\_row(["Modus", modus\_writing])  tabel\_writing.add\_row(["Minimun", min\_writing])  tabel\_writing.add\_row(["Maksimum", max\_writing])  # Menampilkan tabel  print("Tabel nilai tendensi central writing\_score")  print(tabel\_writing) |

Dan Output yang di dapatkan dari program diatas adalah:

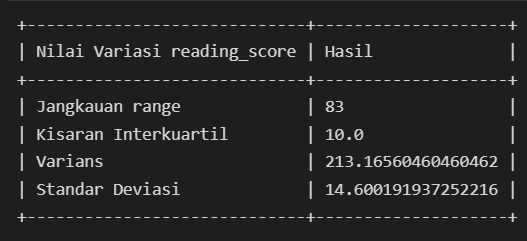


1. **Tentukan nilai Variasi: Jangkauan Range, Kisaran Interkuartil, Varians, Standar Deviasi data *reading\_score* dan *writing\_score.***

Untuk menentukan nilai variasi dari Jangkauan Range, Kisaran Interkuartil, Varians, Standar Deviasi data *reading\_score* dan *writing\_score,* Maka di perlukan bantuan dari libarary python berupa pandas untuk menentukan nilai di tiap data yang akan di visualisasikan serta Prettytable untuk membentuk tabel yang ingin di buat. Adapun program yang di butuhkan untuk membuat nilai variasi dari variabel reading\_score adalah sebagai berikut:

|  |
| --- |
| # Menentukan nilai variasi dari variabel reading\_score  range\_reading = ((df["reading score"].max()) - (df["reading score"].min()))  Q1\_reading = np.quantile(df['reading score'],.25)  Q3\_reading = np.quantile(df['reading score'],.75)  Interquartil\_ReadScore = (Q3\_reading - Q1\_reading)/2  varians\_reading = df['reading score'].var()  Deviasi\_reading =df['reading score'].std()  # Membentuk tabel nilai variasi reading\_score  variasi\_reading = PrettyTable(["Nilai Variasi reading\_score", "Hasil"])  variasi\_reading.align = "l"  # Memuat nilai jangkauan range, kisaran interkuartil, varians, dan Standar deviasi kedalam tabel  variasi\_reading.add\_row(["Jangkauan range", range\_reading])  variasi\_reading.add\_row(["Kisaran Interkuartil", Interquartil\_ReadScore])  variasi\_reading.add\_row(["Varians", varians\_reading])  variasi\_reading.add\_row(["Standar Deviasi", Deviasi\_reading])  # Menampilakan tabel nilai variasi reading\_score  print(variasi\_reading) |

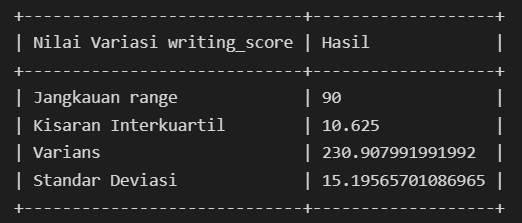
Dan Output yang di dapatkan dari program diatas adalah:



Adapun program yang di butuhkan untuk membuat nilai variasi dari variabel writing\_score adalah sebagai berikut:

|  |
| --- |
| # Menentukan nilai variasi dari variabel writing\_score  range\_writing = ((df["writing score"].max()) - (df["writing score"].min()))  Q1\_writing = np.quantile(df['writing score'],.25)  Q3\_writing = np.quantile(df['writing score'],.75)  Interquartil\_ReadScore = (Q3\_writing - Q1\_writing)/2  varians\_writing = df['writing score'].var()  Deviasi\_writing =df['writing score'].std()  # Membentuk tabel nilai variasi reading\_score  variasi\_writing = PrettyTable(["Nilai Variasi writing\_score", "Hasil"])  variasi\_writing.align = "l"  # Memuat nilai jangkauan range, kisaran interkuartil, varians, dan Standar deviasi kedalam tabel  variasi\_writing.add\_row(["Jangkauan range", range\_writing])  variasi\_writing.add\_row(["Kisaran Interkuartil", Interquartil\_ReadScore])  variasi\_writing.add\_row(["Varians", varians\_writing])  variasi\_writing.add\_row(["Standar Deviasi", Deviasi\_writing])  # Menampilkan tabel nilai variasi reading\_score  print(variasi\_writing) |

Dan Output yang di dapatkan dari program diatas adalah:

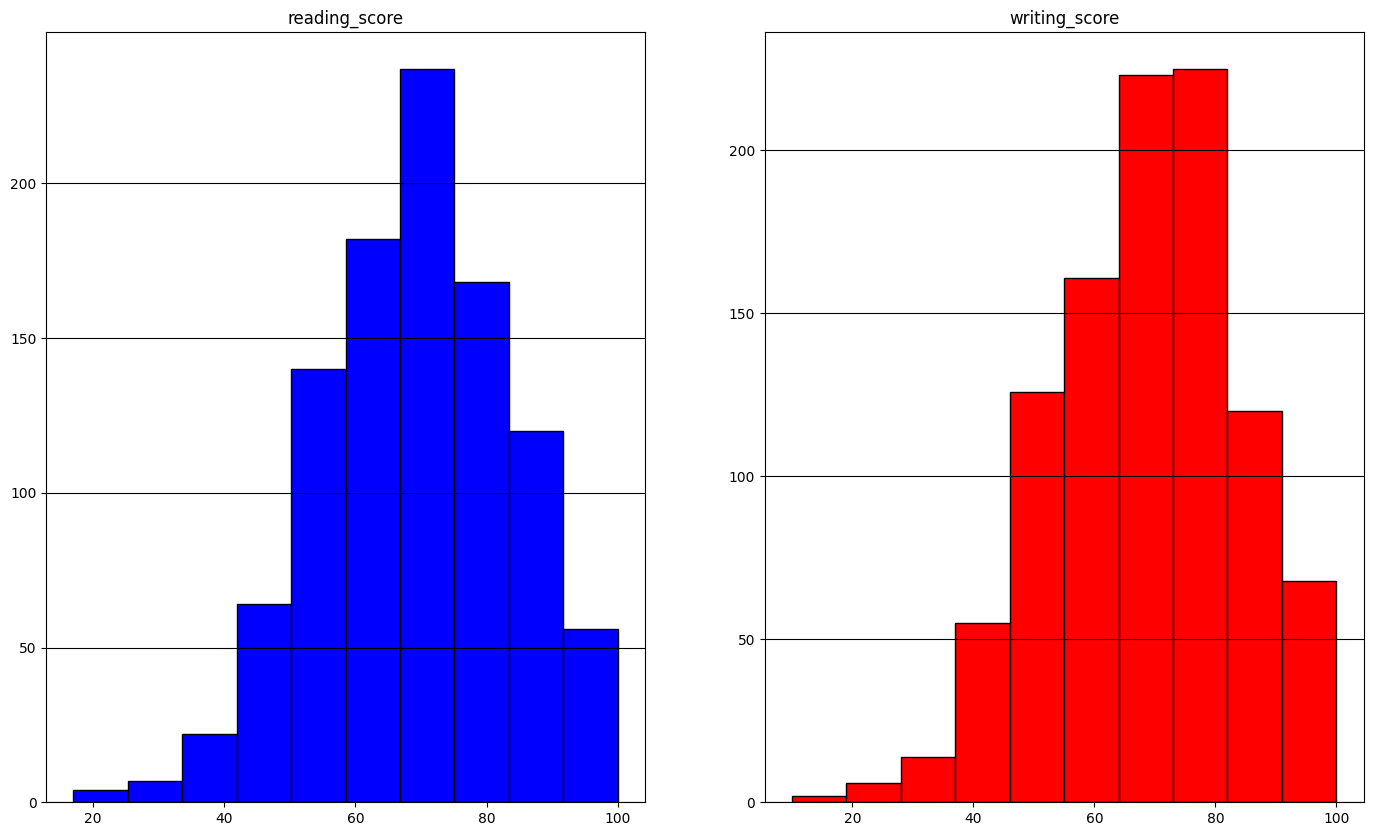


1. **Buat Grafik data Variabel reading\_score dan writing\_score: Histogram, BOX PLOT, dan Tentukan Nilai Skewness dan Kurtosis. Buat interpretasi saudara tentang perbedaan distribusi kedua variable tersebut, terutama pada gambar BOX PLOT dan nilai Skewness - Kurtosisnya.**

Untuk membuat Grafik data Variabel reading\_score dan writing\_score berupa Histogram, BOX PLOT, dan menentukan Nilai Skewness dan Kurtosis serta interpretasi tentang perbedaan distribusi kedua variable tersebut, maka kita perlu melihat grafik serta tabel tersebut, yang mana program di bawah di tujukan untuk menampilkan histogram dari kasus ini, Adapun program yang di perlukan:

|  |
| --- |
| # membuat histogram dengan variabel reading\_score dan writing\_sore  from turtle import color  fig, (histogram\_reading, histogram\_writing) = plt.subplots(1, 2, figsize=(17,10))  histogram\_reading.hist(df['reading score'], color="blue", edgecolor='black')  histogram\_writing.hist(df['writing score'], color="red", edgecolor='black')  #keterangan pada histogram  histogram\_reading.set\_title('reading\_score')  histogram\_reading.grid(axis='y', color='black')  histogram\_writing.set\_title('writing\_score')  histogram\_writing.grid(axis='y', color='black')  # Menampilkan histogram  plt.show() |

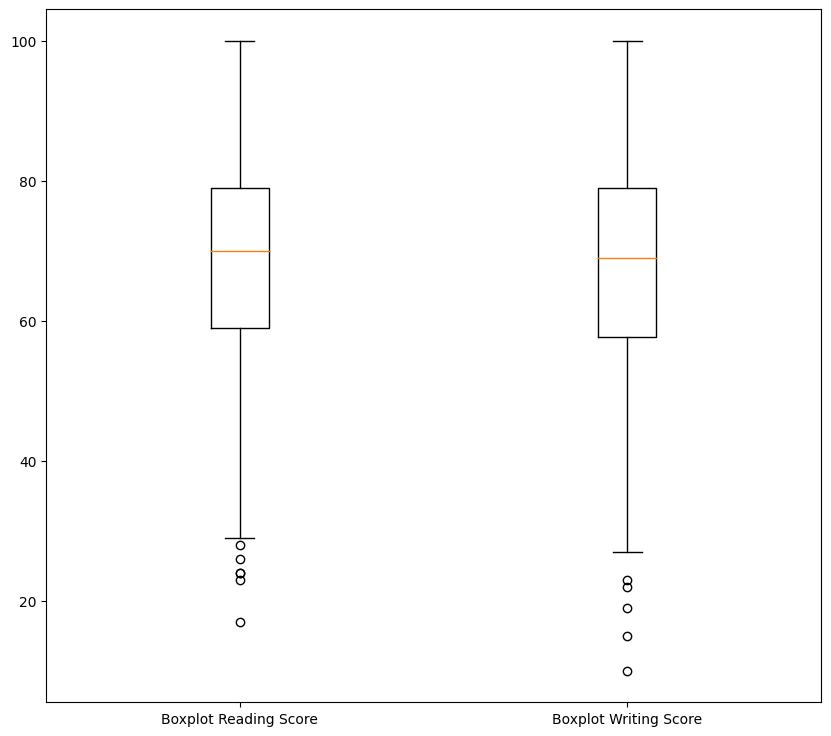
Dan Output yang di dapatkan dari program diatas adalah:



Untuk menampilkan BOX PLOT pada kasus ini, maka program yang di butuh adalah:

|  |
| --- |
| fig, (grafik\_boxplot) = plt.subplots(1, figsize=(10, 9))  #keterangan pada grafik  grafik\_boxplot.boxplot([df['reading score'], df['writing score']])  plt.xticks([1,2],['Boxplot Reading Score','Boxplot Writing Score'])  #menampilkan boxplot  plt.show() |

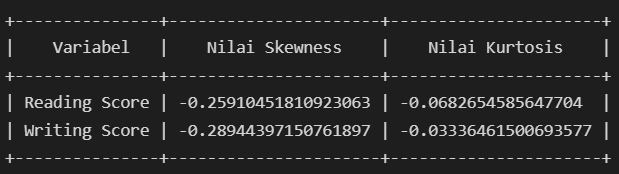
Dan Output yang di dapatkan dari program diatas adalah:



Serta untuk menentukan dan membuat tabel yang memuat nilai skewness dan kurtosis dari variabel reading\_score dan writing\_score, maka di perlukan program berikut:

|  |
| --- |
| # Nilai skewness  skewness\_reading = df['reading score'].skew()  skewness\_writing = df['writing score'].skew()  # Nilai kurtosis  kurtosis\_reading = df['reading score'].kurt()  kurtosis\_writing = df['writing score'].kurt()  # Membuat tabel skewness dan kurtosis  Skewness\_kurtosis = PrettyTable(["Variabel","Nilai Skewness","Nilai Kurtosis"])  Skewness\_kurtosis.add\_row(["Reading Score", skewness\_reading, kurtosis\_reading])  Skewness\_kurtosis.add\_row(["Writing Score", skewness\_writing, kurtosis\_writing])  # Menampilkan tabel nilai skewness dan kurtosis  print(Skewness\_kurtosis) |

Dan Output yang di dapatkan dari program diatas adalah:



Berdasarakan hasil Analisa yang di dapatkan dari grafik histogram dan boxplot, bisa di simpulkan bahwa nilai pada variabel tersebut menjadi sangat berbeda yang bisa di lihat dari pencilan, frekuensi ataupun nilai minimun di grafik.

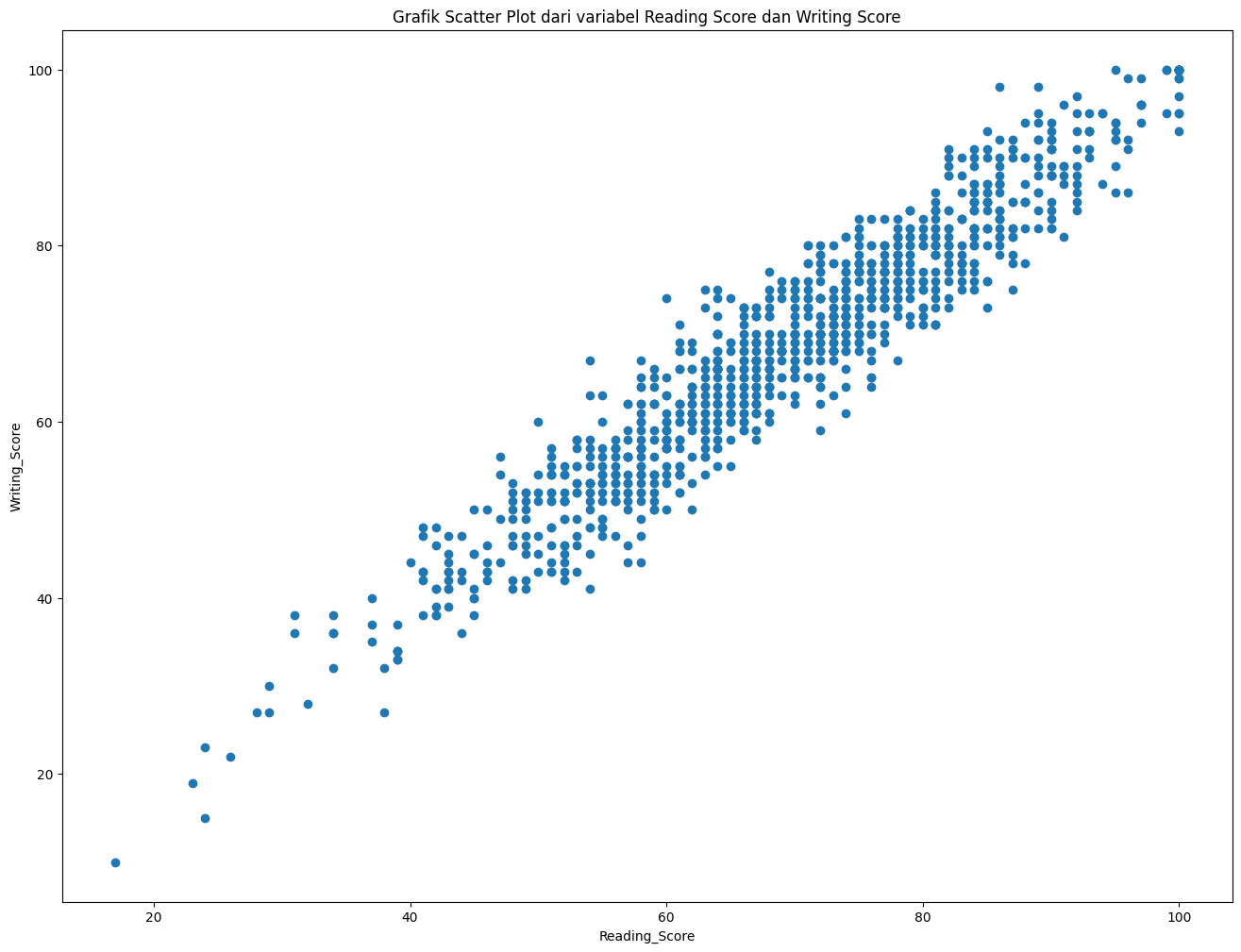
Sedangkan pada nilai skewness dan kurtosis, bisa di lihat sedikit persamaan seperti halnya kedua variabel memiliki nilai negatif di tiap variabel dan sedikit perbedaan seperti halnya angka di belakang koma, yang mana dari fakta tersebut, distribusi setiap data akan relatif sama.

1. **Buat scater plot antara *writing\_score* (dependent variable-Y) versus *reading\_score* (independent variable-X). Tulis interpretasi saudara apakah ada indikasi hubungan linear antar keduanya.**

Untuk membuat scater plot antara writing\_scoren dan reading\_score, maka di perlukan program sebagai berikut:

|  |
| --- |
| #Membuat grafik scater plot antara writing\_score versus reading\_score.  #Menentukan ukuran grafik  plt.figure(figsize=(16,12))  plt.scatter(df["reading score"], df["writing score"])  # Penamaan dan penetuan posisi x dan y  plt.xlabel('Reading\_Score')  plt.ylabel('Writing\_Score')  #Memberi nama grafik dan menampilkan grafik  plt.title('Grafik Scatter Plot dari variabel Reading Score dan Writing Score')  plt.show() |

Dan Output yang di dapatkan dari program diatas adalah:

****

Pada grafik diatas dapat di simpulkan bahwa antara kedua variabel akan dapat memiliki kecenderungan dimana aka nada kemungkinan saat salah satu nilai mengalami kenaikan, maka nilai variabel lainya juka memiliki kesempatan mengalami kenaikan.

Demikian jawaban dari beberapa kasus yang di berikan sebelumnya, terkait program, dataset dan hal penting lainya akan tercantum pada link di bawah:

[TUGAS\_II\_PROBAS\_DAN\_STATISTIKA](https://github.com/Alfinorihsan/TUGAS_II_PROBAS_DAN_STATISTIKA-)