**LAPORAN UTS PROBALITAS DAN STATISTIKA**

**Nama : Muhammad Novan Saputra**

**Nim : 2109106032**

**Kelas : Informatika A’21**

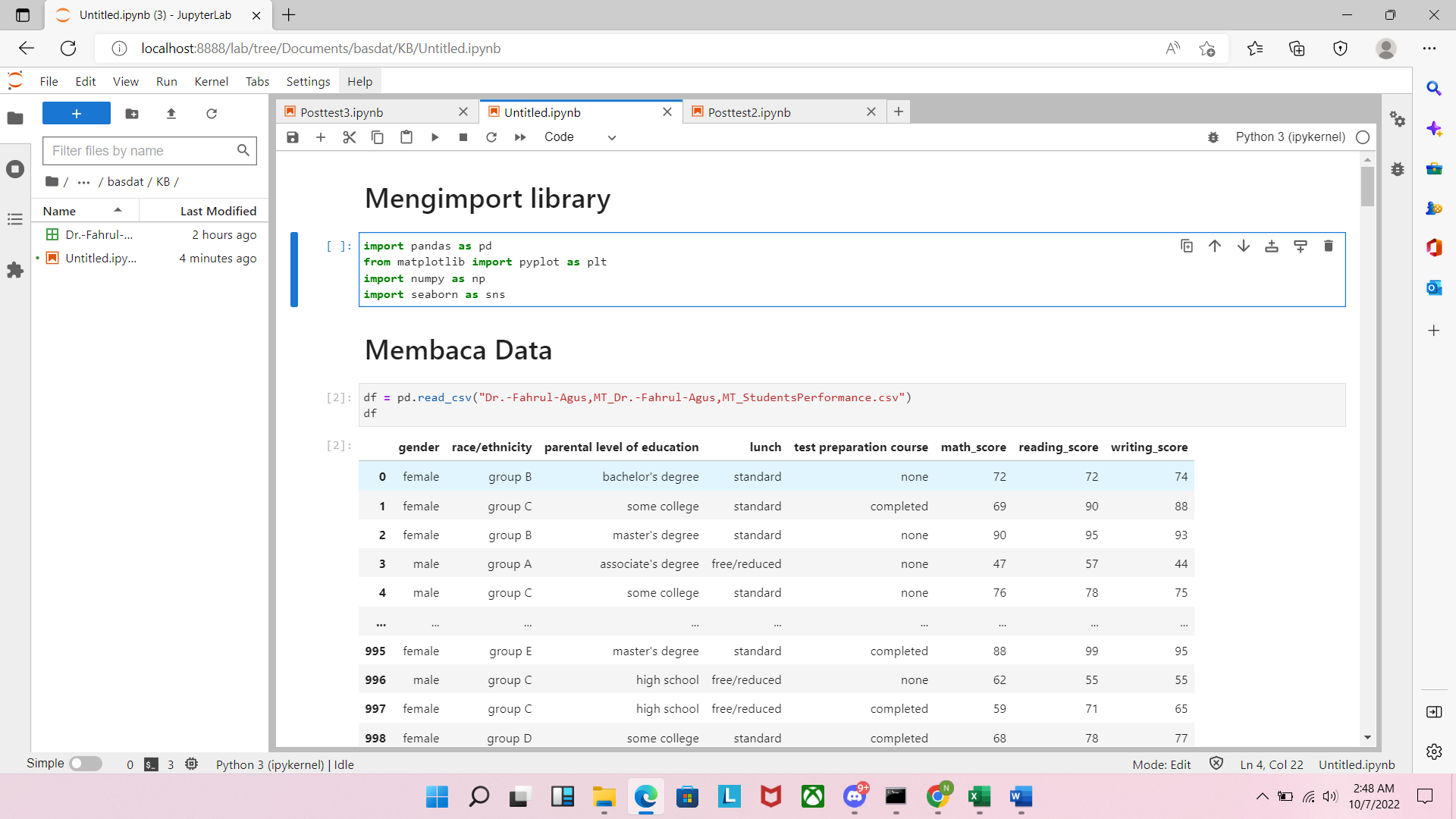
**Matkul : Probalitas Dan Statistika**

1. **Pendahuluan**

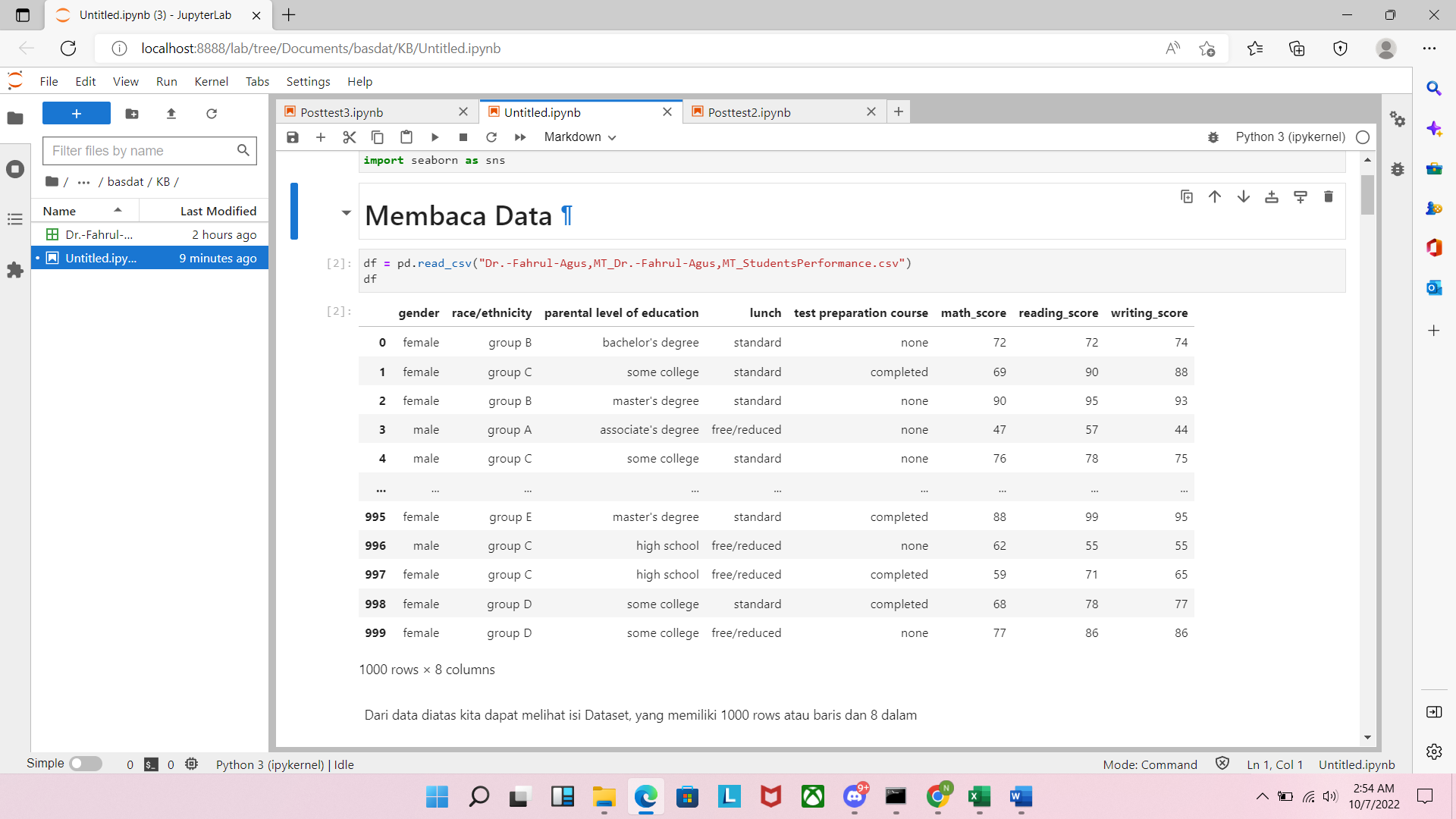
Program ini menggunakan library pandas, numpy, matplotlib, dan seaborn. Program yang dibuat ini untuk menyelesaikan analisis yang telah diberikan di mols. Analisis yang dilakukan yaitu menganalisis dataset yang telah diberikan. Berikut adalah analisis yang telah di tentukan :

1. Buat Tabel Distribusi Frekuensi dan Histogram variable *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 variable tersebut.
3. Tentukan nilai variasi : Jangkauan Range, kisaran interkuartil varians, standar Deviasi data *reading\_score* dan *writing\_score.*
4. Buat Grafik data variable *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 Scatter plot antara *writing\_score* (dependent variable y) versus *reading\_score* (independent variable x). Tulis interpretasi saudara apakah ada indikasi hubungan linier.
6. **Penjelasan**

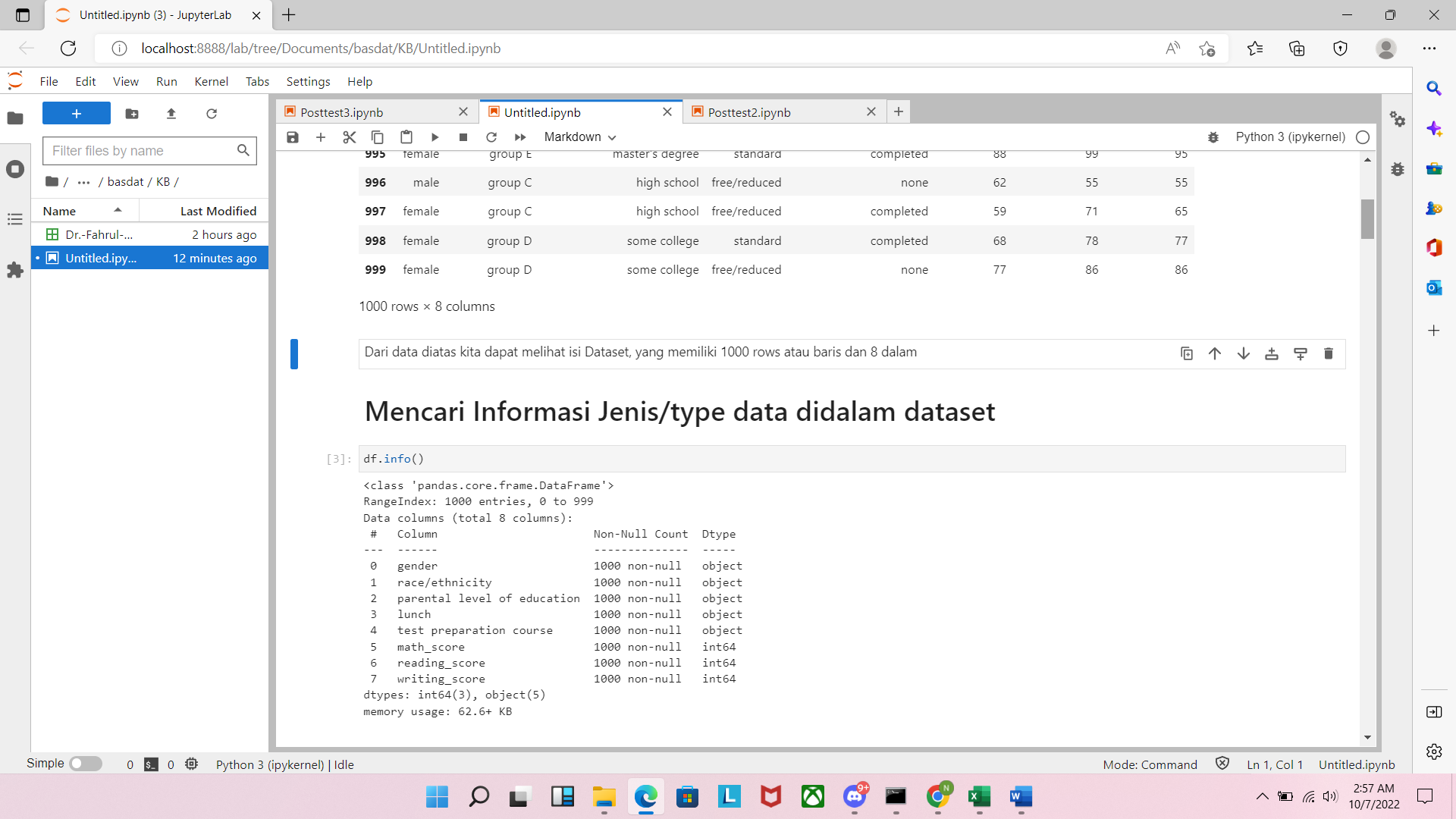
Berikut ini cara memasuk library



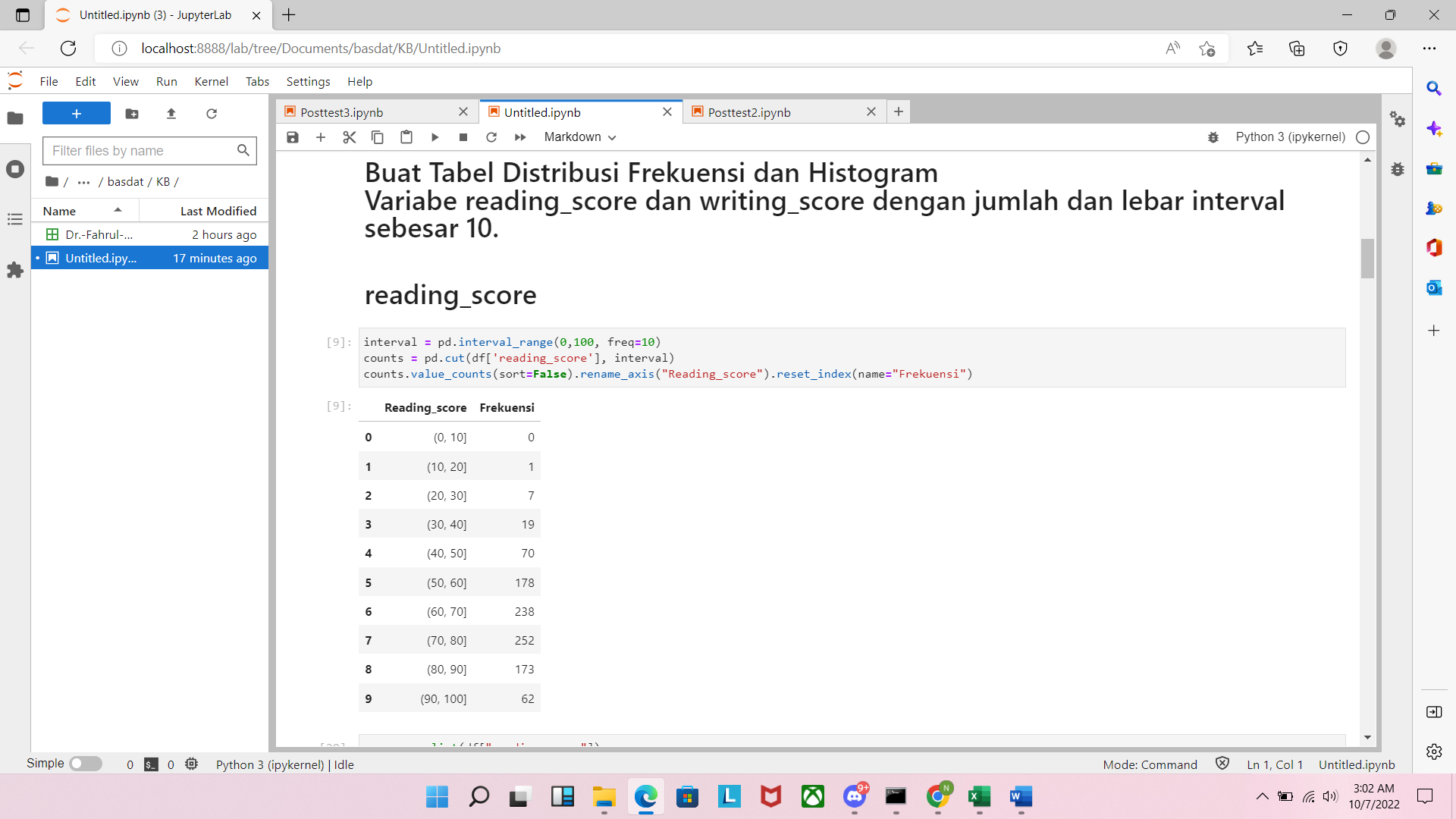
Kita dapat membaca atau melihat Dataset dengan cara seperti dibawah



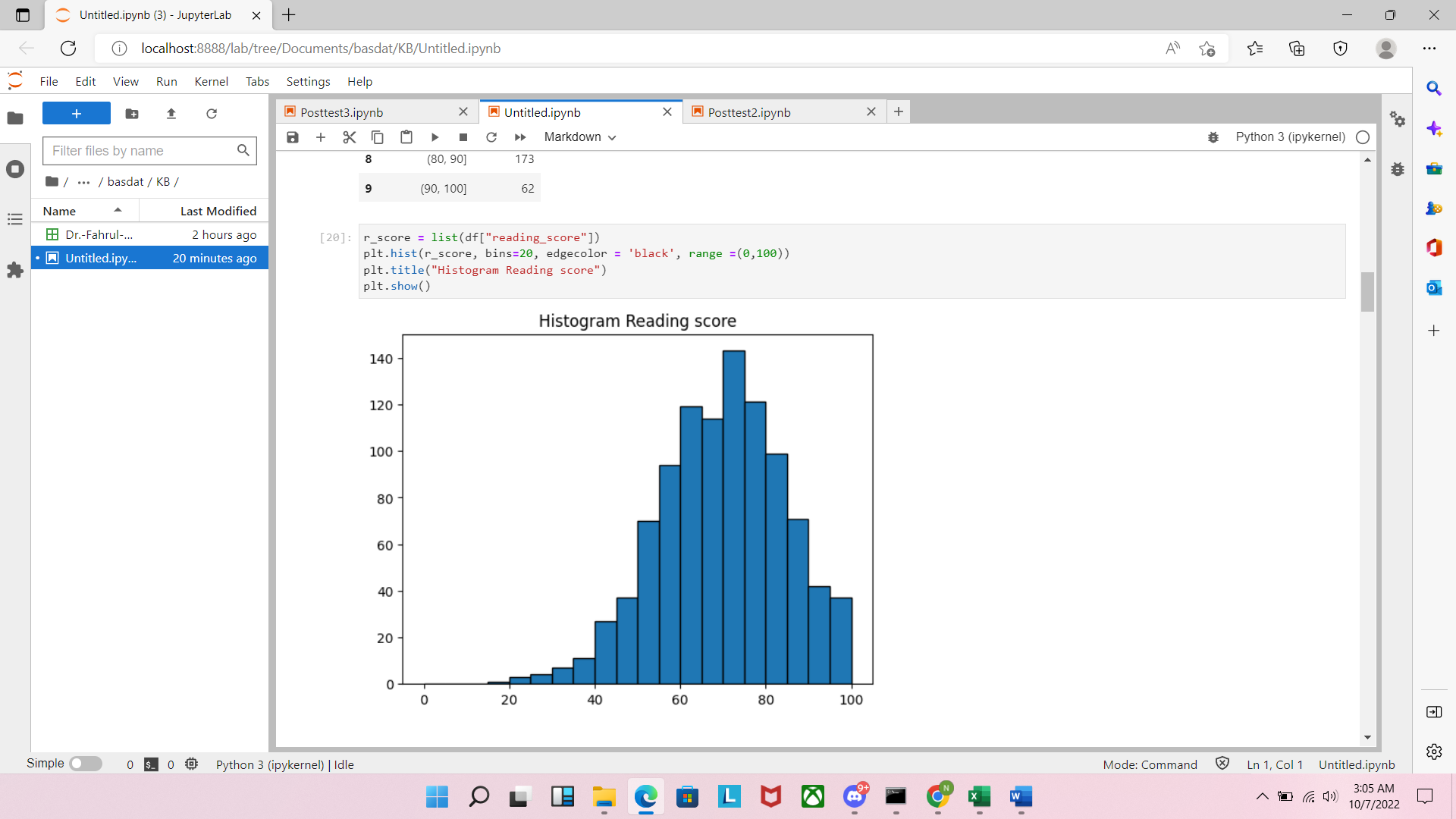
Mancari tau jenis/type data dari setiap variable yang ada di data set tersebut dengan df.info()

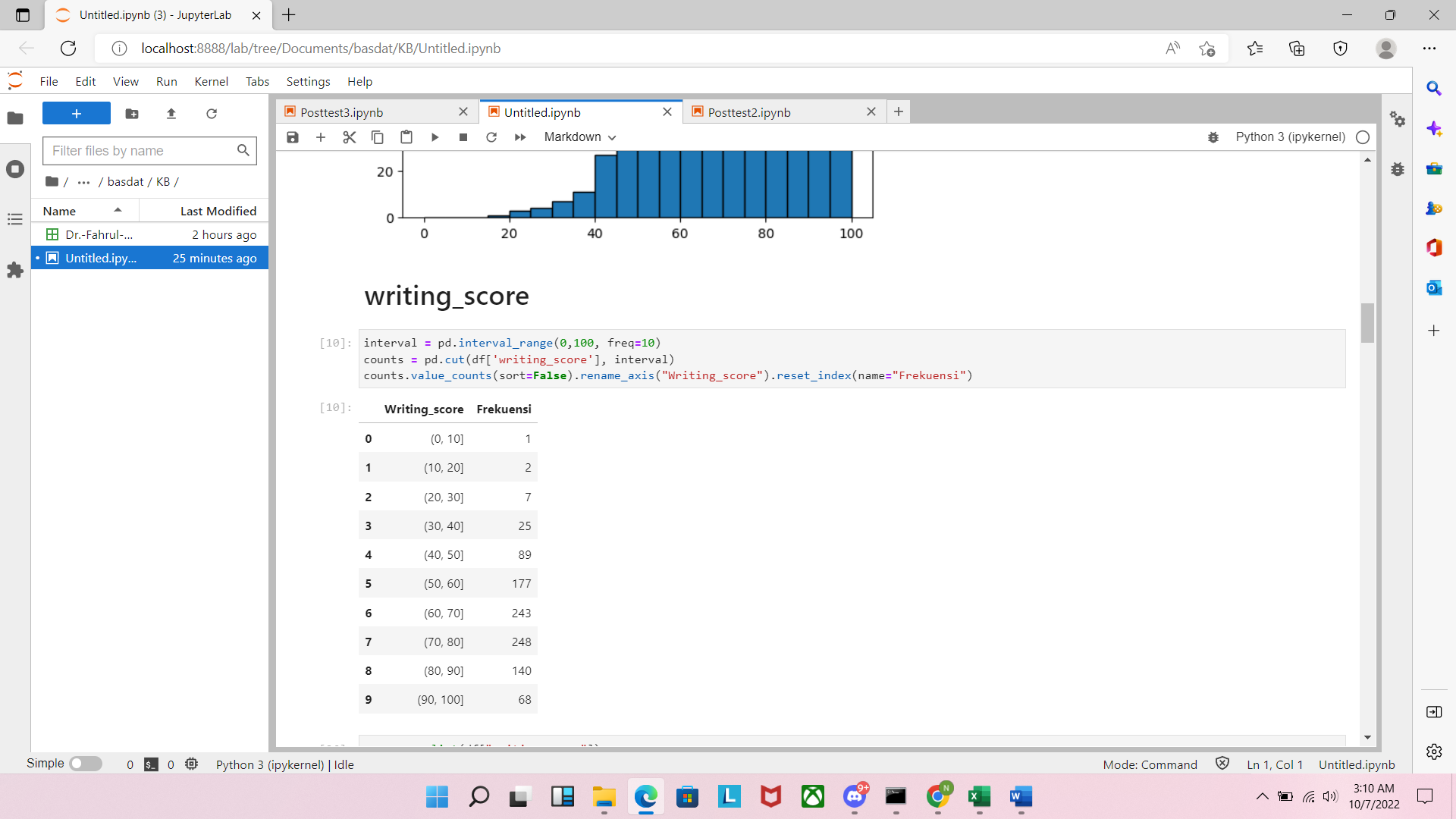


Buat Tabel Distribusi Frekuensi dan Histogram variable *reading\_score* dan *writing\_score* dengan jumlah dan lebar interval sebesar 10.

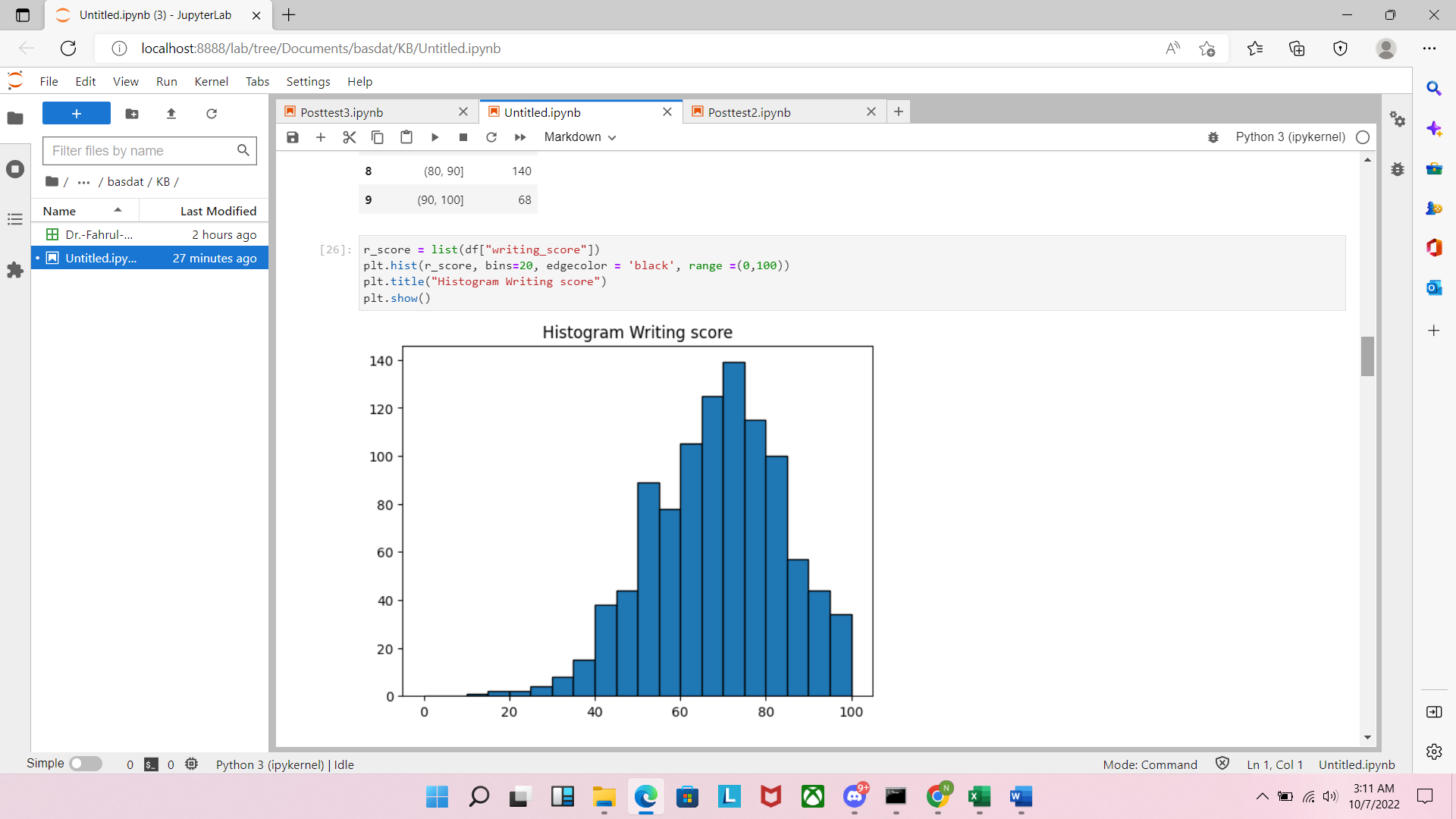


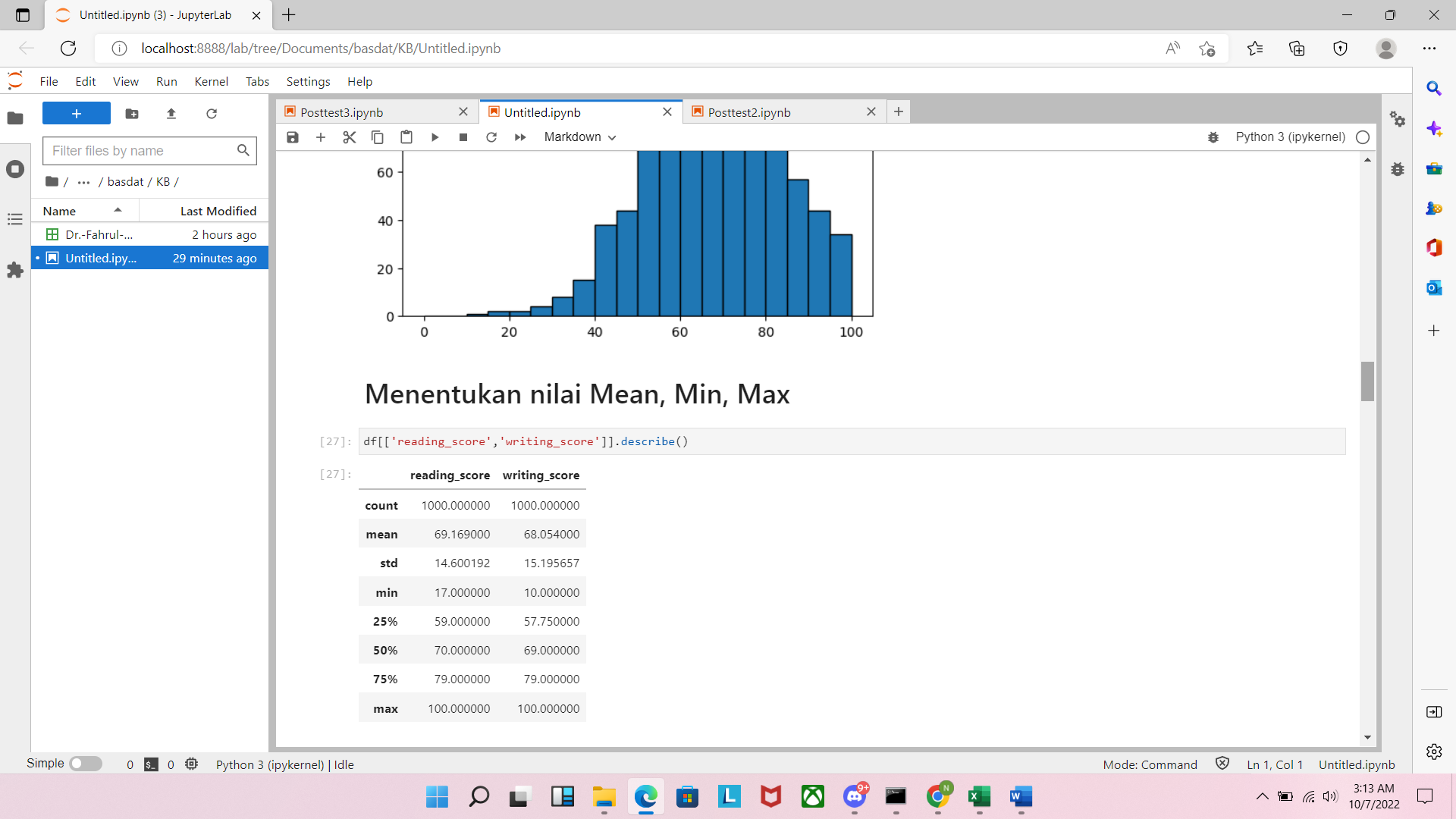
Gambar diatas merupakan tabel distribusi Frekuensi Reading\_sorce, dan dibawah merupakan Histrogram Reading data

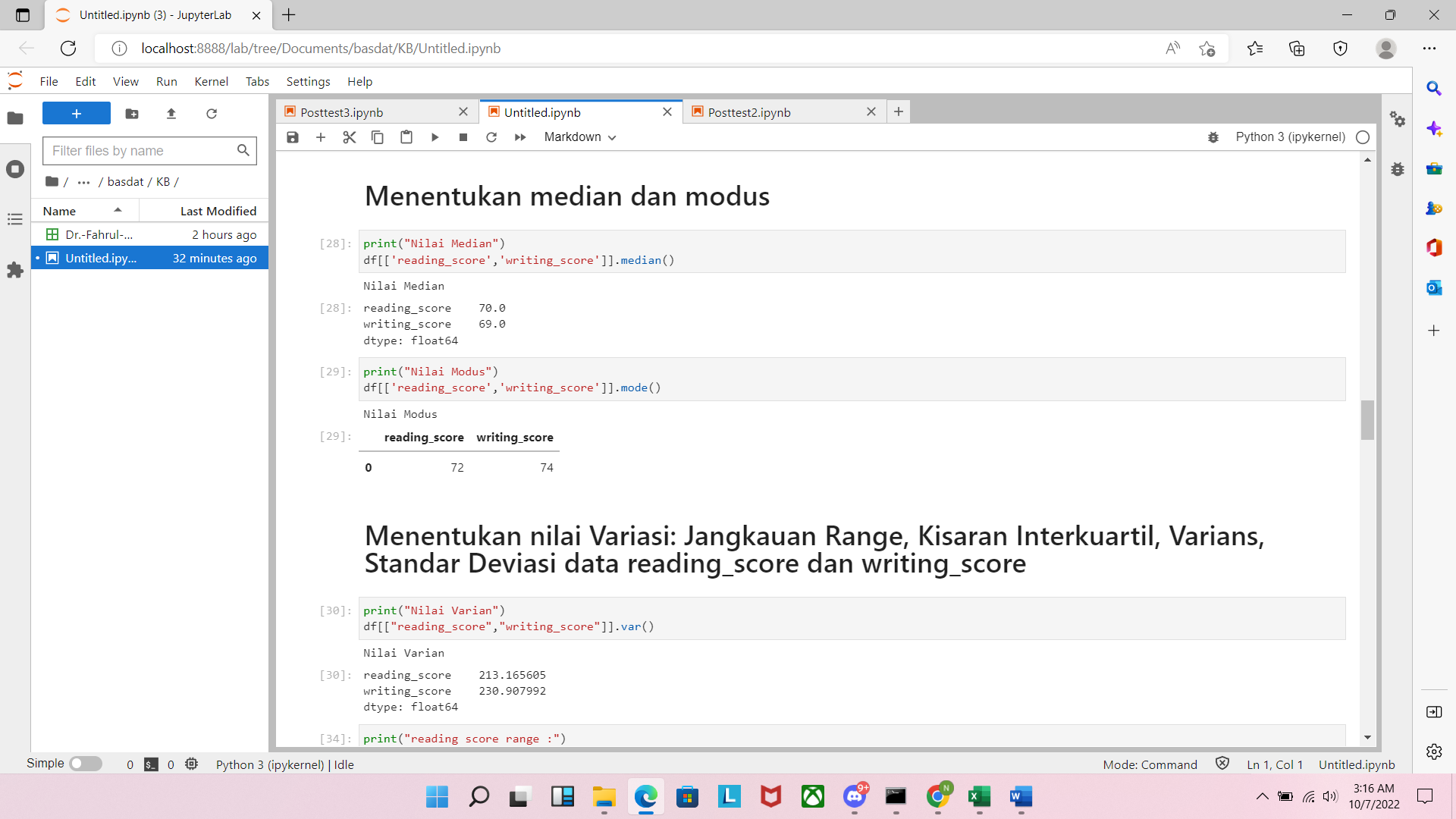




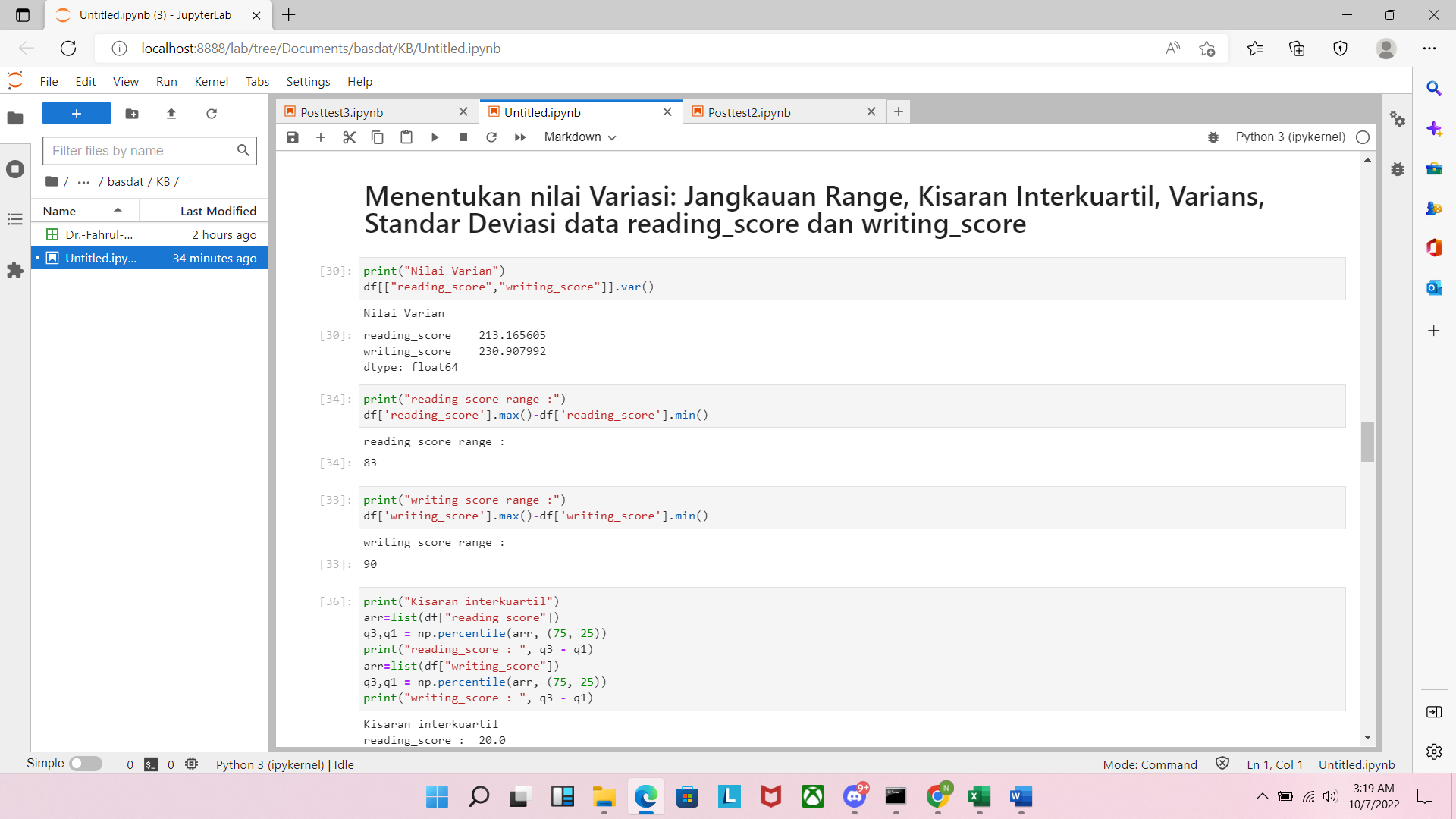
Gambar diatas merupakan tabel distribusi Frekuensi Writing\_sorce, dan dibawah merupakan Histrogram Writing data

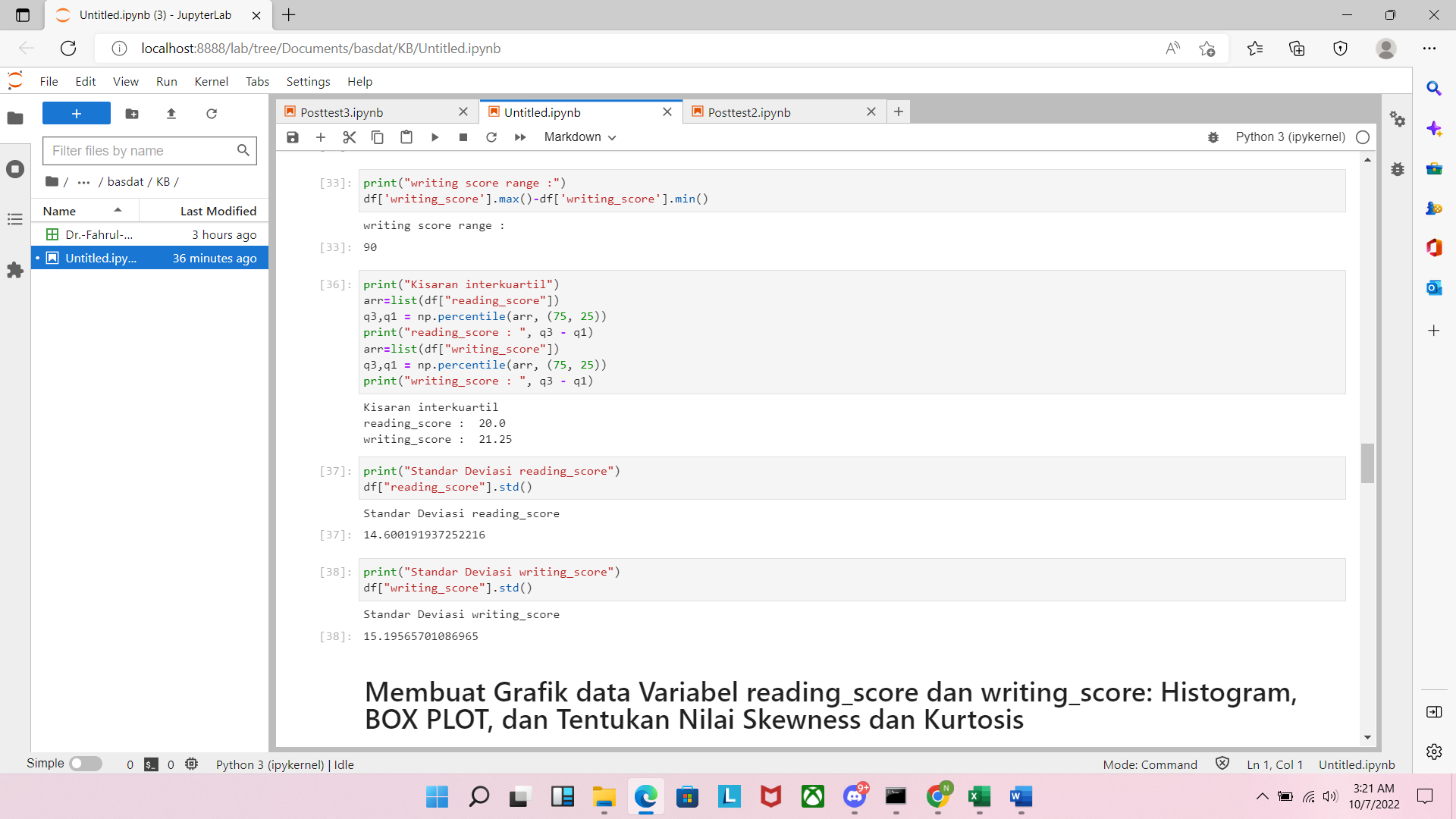
Tentukan nilai tendensi central : Rata-rata, Median, Modus, Min dan Max data kedua variable tersebut.



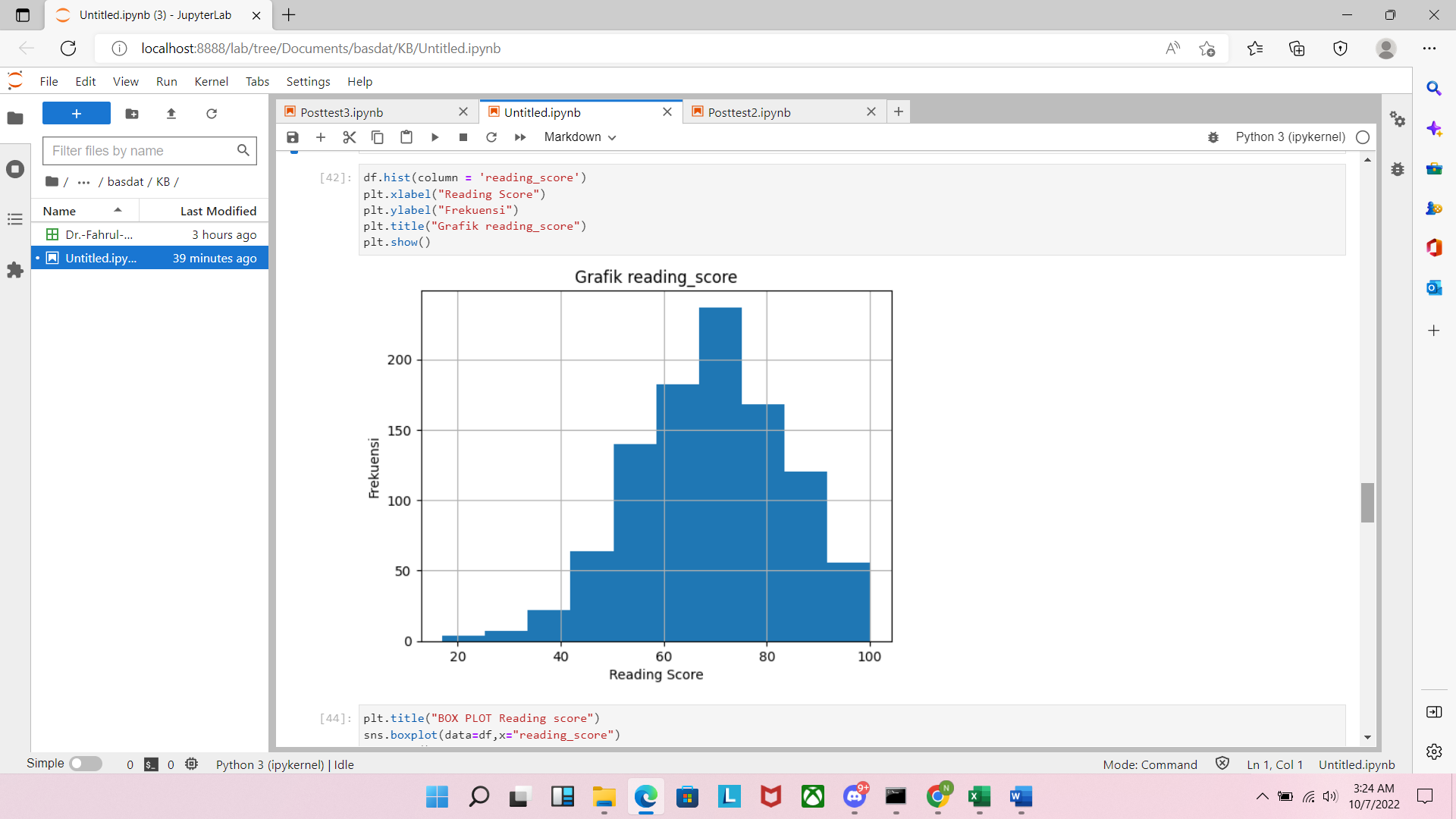


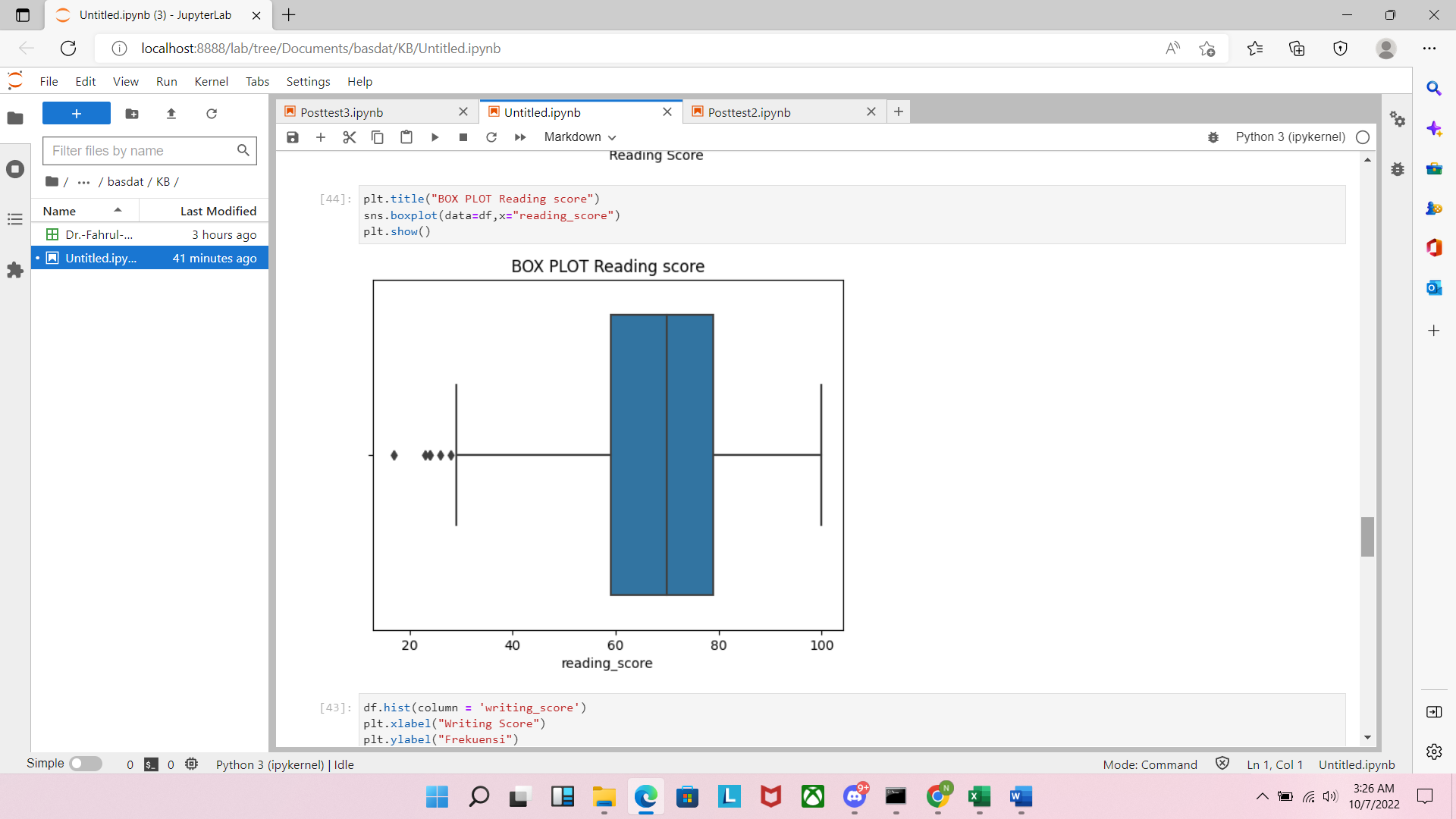
Tentukan nilai variasi : Jangkauan Range, kisaran interkuartil varians, standar Deviasi data *reading\_score* dan *writing\_score.*

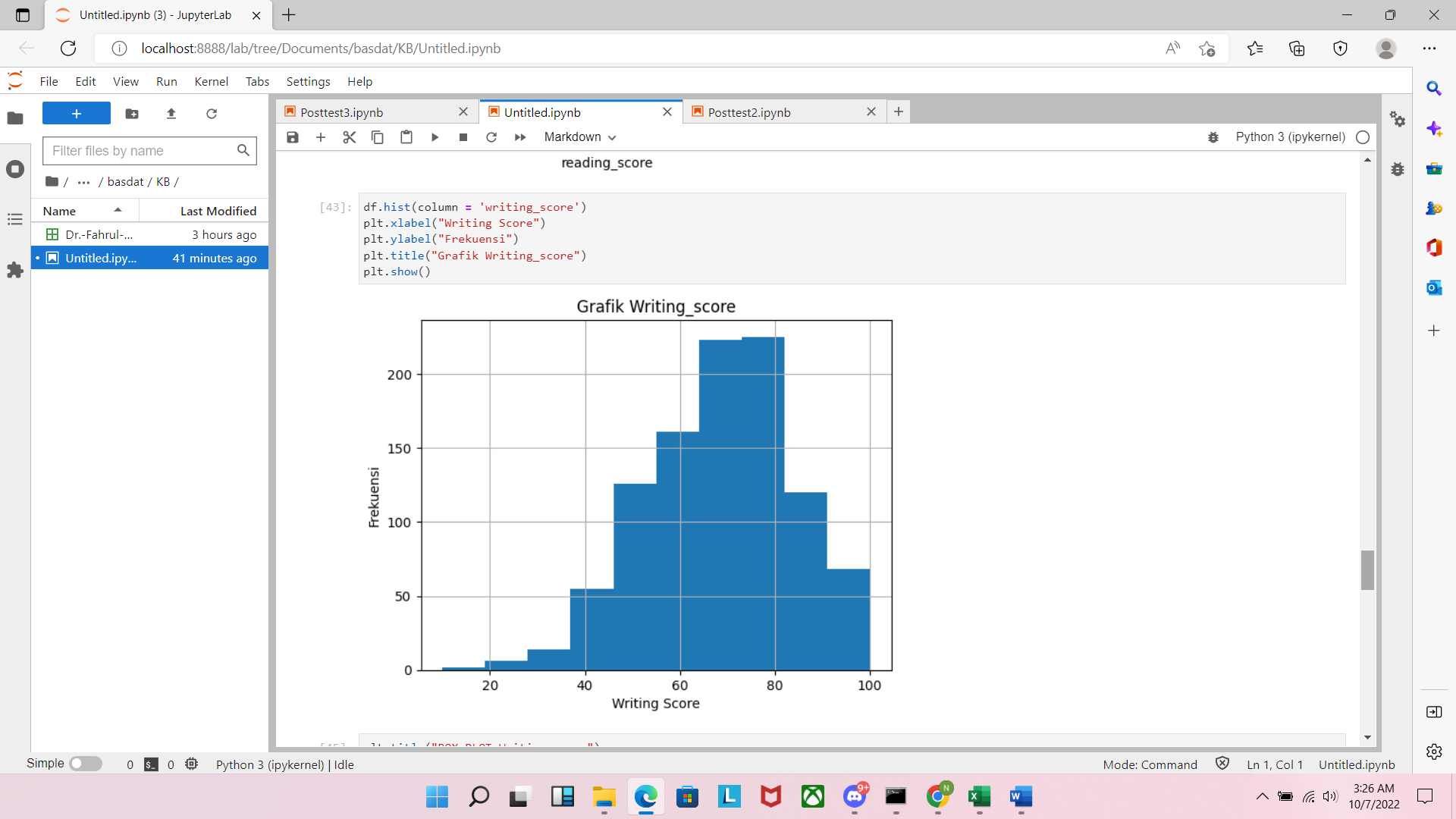


Buat Grafik data variable *reading\_score* dan *writing\_score*; histogram, box plot, dan tentukan nilai

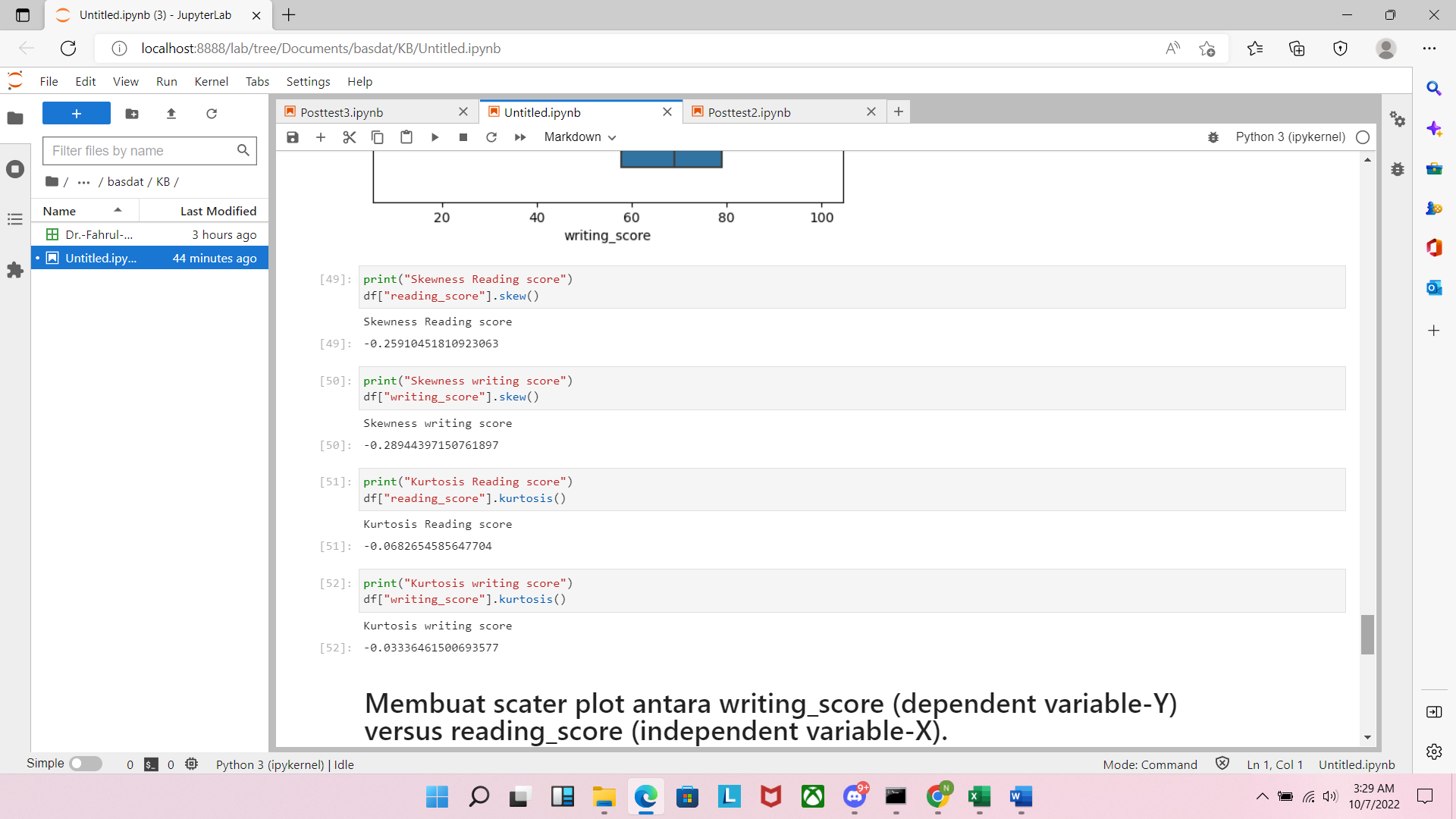
skewness dan kurtosis. Buat interpretasi saudara tentang perbedaan distribusi kedua variable tersebut, terutama pada gambar BOX plot dan nilai skewness-kurtosisnya.







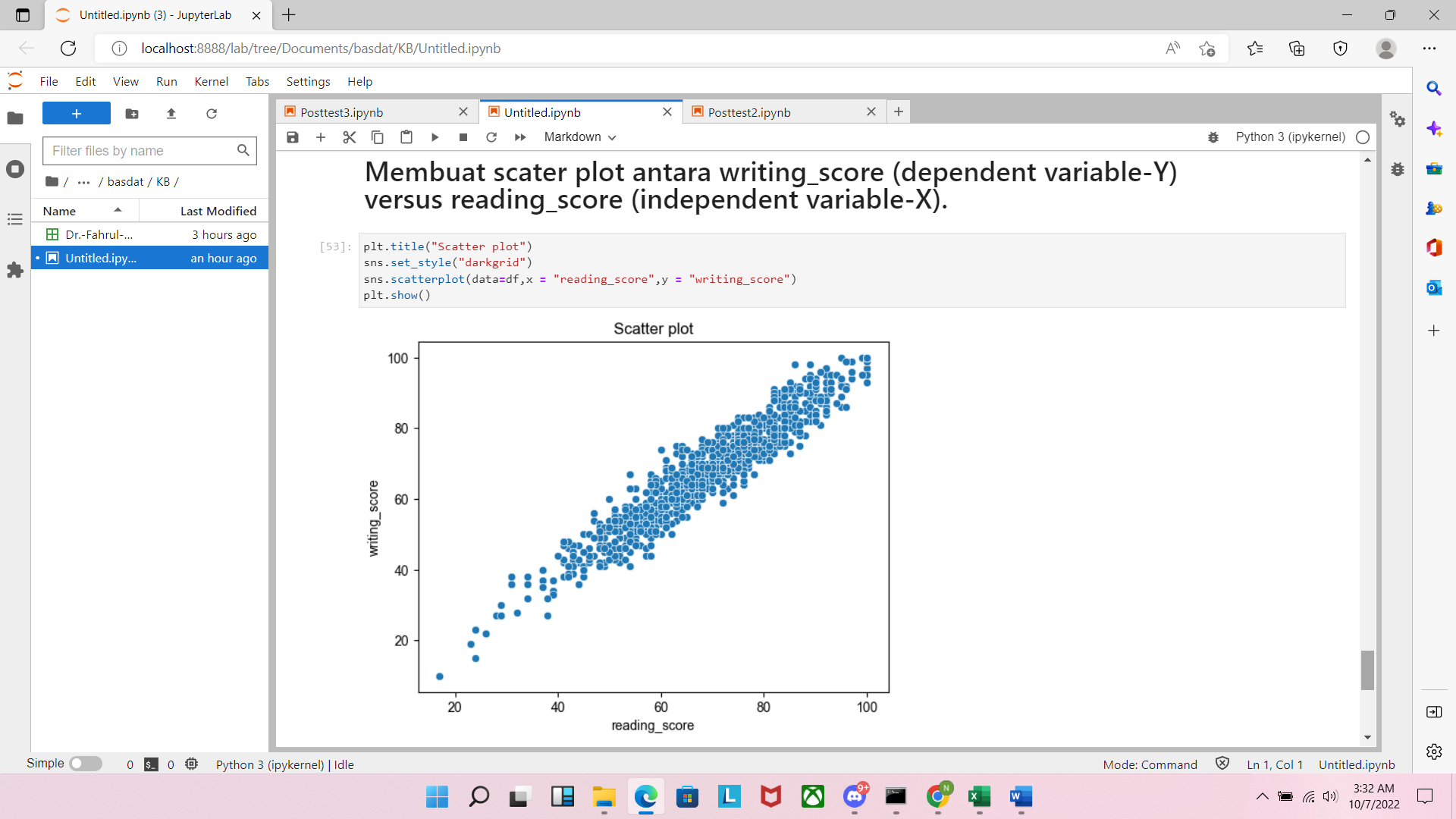




Dari grafik data diatas dapat dilihat bahwa terdapat perbedaan pada outliernya yang dimana outlier dari *reading\_score* cenderung berhimpit antar nilainya, sedangkan outlier dari *writing\_score*  lebih menyebar untuk nilai nilainya.

Dari nilai skewness dan kurtosisnya dapat disimpulkan bahwa persebaran data dari atribut *reading\_score* dan *writing\_score*  dapat dikatakan normal nilai skewness dari 2 atribut tersebut mendekati 0 tetapi kurtosisi dari attribute tersebut cenderung runcing atau dengan kata lain cenderung mengelompok (homogen) dikarenakan nilai kurtosisnya yang jauh dari nilairefrensi kurtosis yaitu 3.

Buat Scatter plot antara *writing\_score* (dependent variable y) versus *reading\_score* (independent variable x). Tulis interpretasi saudara apakah ada indikasi hubungan linier.



Indikasi hubungan linier antara atribut *reading\_score* dengan *writing\_score* adalah berupa korelasi, yang dimana mode linier satu variable berubah karena terkait dengan perubahan pada variable lain secara teratur dengan arah yang sama (korelasi linier positif).