


Rencana Tugas Mandiri ke-3
Mata Kuliah Analisis Data Eksploratif
Sub-CMPK-3 Statistika Deskriptif dan Grouping Dataset

| | | | | | |
|---|---|--|---|----------|---|
|  | Universitas Pembangunan Nasional Veteran Jawa Timur | | | | |
| | Fakultas Ilmu Komputer | | | | |
| | Program Studi Sains Data | | | | |
| Rencana Tugas Mahasiswa | | | | | |
| Mata Kuliah | Analisis Data Eksploratif | | | | |
| Kode | SD211123 | SKS | 3 | Semester | 5 |
| Dosen Pengampu | Tresna Maulana Fahrudin, S.ST., M.T. | | | | |
| Bentuk Tugas | Studi Kasus | | | | |
| Judul Tugas | Tugas 3: Statistika Deskriptif dan Grouping Dataset | | | | |
| Sub CPMK | Sub-CMPK-3 Statistika Deskriptif dan Grouping Dataset | | | | |
| Deskripsi Tugas | Tugas ini bertujuan agar mahasiswa mampu untuk melakukan <ol style="list-style-type: none">Statistika Deskriptif: Distribusi Uniform, Distribusi Normal, Skewness dan KurtosisGrouping Dataset: Pivoting Tabel | | | | |
| Metode Pengerjaan Tugas | 1. | Mencari referensi dan menyelesaikan permasalahan studi kasus yang berkaitan dengan statistika deskriptif dan grouping dataset | | | |
| | 2. | Tugas dikerjakan secara individu | | | |
| | 3. | Tugas diupload di <i>e-learning</i> (masing-masing mahasiswa). Tugas dikumpulkan paling lambat pada H-1 (maksimal 23.59 WIB) | | | |
| | 4. | Tugas diketik pada kertas A4, <i>font</i> ukuran 12, Times New Roman, dan spasi 1. Tidak ada ketentuan jumlah halaman halaman. Tuliskan nama mahasiswa, NPM, dan kelas paralel | | | |
| Bentuk dan Format Luaran | Hasil dokumentasi penyelesaian tugas dengan sistematika dan format yang telah ditentukan | | | | |
| Indikator, Kriteria dan Bobot Penelitian | | | | | |
| Indikator: Ketepatan dalam mencari solusi/teknik/pendekatan untuk menyelesaikan permasalahan berkaitan dengan statistika deskriptif dan grouping dataset Kriteria: Non-test Bobot: 4% Rubrik: Holistik | | | | | |
| Lain-Lain | | | | | |
| Daftar Pustaka: S. Kumar Mukhiya., U. Ahmed..2020. <i>Hands-on Exploratory Data Analysis with Python</i> . Birmingham: Packt Publishing | | | | | |

Implementasikan teknik-teknik statistika deskriptif dan pivoting tabel untuk menyelesaikan berbagai permasalahan studi kasus pada dataset.

1. Statistika Deskriptif (Distribusi Uniform, Distribusi Normal, Skewness, dan Kurtosis)

a. Distribusi Uniform

| | | | | |
|-----|-----|-----|-----|-----|
| 3,5 | 1,5 | 8,2 | 2,6 | 4,5 |
| 9,1 | 2,7 | 6,8 | 4,3 | 3,5 |
| 3,6 | 7,4 | 7,3 | 8,9 | 7,1 |
| 2,1 | 9,5 | 2,8 | 2,3 | 8,4 |
| 8,6 | 12 | 2,1 | 3,5 | 1,5 |
| 6,7 | 5,5 | 6,7 | 2,6 | 5,5 |
| 6,6 | 6,9 | 8,3 | 8,9 | 7,5 |
| 4,3 | 9,7 | 2,6 | 3,3 | 8,8 |

Catatan: **ph.csv (dataset)**

Berapa persentase probabilitas $x < 5$ berdasarkan persamaan dalam distribusi uniform dengan menggunakan kode *script* pemrograman Python secara *scratch*

- Carilah nilai a dan b terlebih dahulu
- Gunakan persamaan *mean* dan *standard deviation* sebagai bekal untuk mendapatkan nilai a dan b sesuai dengan sifat dasar statistika dalam distribusi uniform berikut

| | |
|---------------|-----------------------------------|
| Mean | $(\alpha + \beta)/2$ |
| Median | $(\alpha + \beta)/2$ |
| Mode | any $x, \alpha \leq x \leq \beta$ |
| Variance | $(\beta - \alpha)^2/12$ |
| Std deviation | $(\beta - \alpha)/\sqrt{12}$ |
| Skewness | 0 |
| Kurtosis | -1.2 |

- Selanjutnya hitung probabilitas distribusi uniform berdasarkan nilai a dan b yang sudah didapatkan menggunakan persamaan berikut

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} \frac{1}{b-a} dx$$

- Visualisasikan dataset menggunakan histogram, periksalah apakah dataset berdistribusi uniform atau tidak. Jelaskan beserta alasannya.

Sumber: <https://www.real-statistics.com/other-key-distributions/uniform-distribution/>

b. Distribusi Normal

| x | f(x) |
|-------|------|
| -5,84 | |
| -4,86 | |
| -3,96 | |
| -3,74 | |
| -2,86 | |
| -2,35 | |
| -2,25 | |
| -2,16 | |
| -1,8 | |
| -1,69 | |
| -1,53 | |
| -1,52 | |
| -1,17 | |
| -1,17 | |
| -1,12 | |
| -0,94 | |
| -0,89 | |
| -0,86 | |
| -0,73 | |
| -0,69 | |
| -0,44 | |
| -0,37 | |
| -0,28 | |
| -0,2 | |
| -0,15 | |
| -0,09 | |
| 0,12 | |
| 0,14 | |
| 0,31 | |
| 0,33 | |
| 0,4 | |
| 0,56 | |
| 0,73 | |
| 0,73 | |
| 0,91 | |
| 0,92 | |

| | |
|------|--|
| 1,2 | |
| 1,21 | |
| 1,55 | |
| 1,8 | |
| 1,9 | |
| 1,91 | |
| 1,91 | |
| 1,93 | |
| 2,14 | |
| 2,45 | |
| 3 | |
| 3,13 | |
| 3,54 | |
| 4,12 | |
| 4,85 | |
| 5,96 | |

Catatan: **normal_distribution.csv (dataset)**

- Hitunglah persamaan $f(x)$ dari nilai x di atas berdasarkan persamaan dalam distribusi normal dengan menggunakan kode *script* pemrograman Python secara *scratch*.
- Gunakan persamaan distribusi normal sebagai berikut:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- Visualisasikan antara nilai x dan $f(x)$ menggunakan diagram garis, periksalah apakah dataset berdistribusi normal atau tidak. Jelaskan beserta alasannya.

c. Skewness

| x | f(x) |
|------|------|
| -10 | |
| -3,3 | |
| -3,2 | |
| -3,1 | |
| -3 | |
| -2,9 | |
| -2,8 | |
| -2,7 | |
| -2,6 | |
| -2,5 | |
| -2,4 | |
| -2,3 | |
| -2,2 | |
| -2,1 | |
| -2 | |
| -1,9 | |
| -1,8 | |
| -1,7 | |
| -1,6 | |
| -1,5 | |
| -1,4 | |
| -1,3 | |
| -1,2 | |
| -1,1 | |
| -1 | |
| 0 | |
| 1 | |
| 1,1 | |
| 1,2 | |
| 1,3 | |
| 1,4 | |
| 1,5 | |
| 1,6 | |
| 1,7 | |
| 1,8 | |

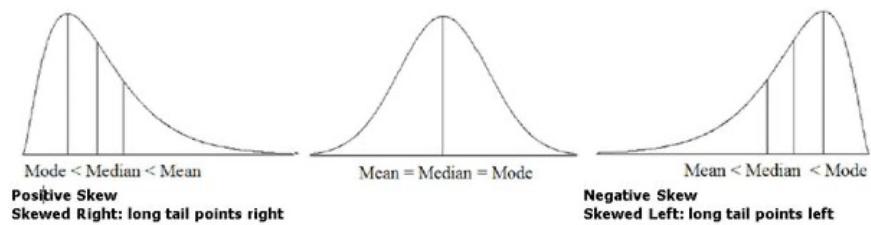
| | |
|-----|--|
| 1,9 | |
| 2 | |
| 2,1 | |
| 2,2 | |
| 2,3 | |
| 2,4 | |
| 2,5 | |
| 2,6 | |
| 2,7 | |
| 2,8 | |
| 2,9 | |
| 3 | |
| 3,1 | |
| 3,2 | |
| 3,3 | |
| 3,4 | |
| 5 | |

Catatan: **skewness.csv (dataset)**

- Hitunglah persamaan $f(x)$ dari nilai x di atas berdasarkan persamaan dalam **distribusi normal** dengan menggunakan kode *script* pemrograman Python secara *scratch*.
- Visualisasikan antara nilai x dan $f(x)$ menggunakan diagram garis, periksalah apakah dataset berdistribusi normal atau tidak. Jelaskan beserta alasannya.
- Selanjutnya, gunakan persamaan *skewness* sebagai berikut (ke Python):

$$\text{Skewness} = \frac{3 (\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

Apakah ada kecondongan distribusi data, ke arah positif atau ke arah negatif ? Jelaskan beserta alasannya.



d. Kurtosis

| x | f(x) |
|-------|------|
| -29,4 | |
| -29,3 | |
| -29,2 | |
| -29,1 | |
| -2,5 | |
| -2,4 | |
| -2,3 | |
| -2,2 | |
| -2,1 | |
| -2 | |
| 0 | |
| 1 | |
| 1,1 | |
| 1,2 | |
| 1,3 | |
| 1,4 | |
| 1,5 | |
| 1,6 | |
| 1,7 | |
| 1,8 | |
| 1,9 | |
| 2 | |
| 2,1 | |
| 2,2 | |
| 2,3 | |
| 2,4 | |
| 2,5 | |
| 2,6 | |
| 2,7 | |
| 2,8 | |
| 2,9 | |
| 3 | |
| 3,1 | |
| 3,2 | |

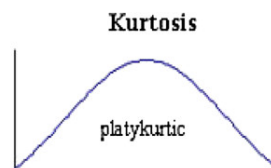
| | |
|------|--|
| 3,3 | |
| 3,4 | |
| 3,5 | |
| 3,6 | |
| 3,7 | |
| 3,8 | |
| 31,1 | |
| 31,2 | |
| 31,3 | |

Catatan: kurtosis.csv (dataset)

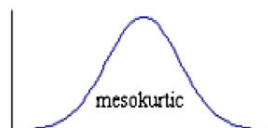
- Hitunglah persamaan $f(x)$ dari nilai x di atas berdasarkan persamaan dalam **distribusi normal** dengan menggunakan kode *script* pemrograman Python secara *scratch*.
- Visualisasikan antara nilai x dan $f(x)$ menggunakan diagram garis, periksalah apakah dataset berdistribusi normal atau tidak. Jelaskan beserta alasannya.
- Selanjutnya, gunakan persamaan *kurtosis* sebagai berikut (ke Python):

$$\alpha_4 = \frac{\frac{1}{n} \sum (X - \bar{X})^4}{s^4}$$

Distribusi data mengarah ke platykurtic, mesokurtic, atau leptokurtic? Jelaskan beserta alasannya.



Platykurtic, kurtosis < 3, the tails are very thin compared to the normal distribution



Mesokurtic, kurtosis = 3



Leptokurtic, kurtosis > 3, fat tails indicate that the distribution produces more outlier

2. Grouping Dataset: Pivoting Table

| TANGGAL | BARANG | SUPLIER | BULAN | TAHUN | KUARTAL | JUMLAH | HARGA | TOTAL |
|----------|--------------|---------|----------|-------|-----------|--------|--------------|-----------------|
| 01/01/15 | Cabai | Bejo | January | 2015 | Kuartal 1 | 55 | Rp 30.550 | Rp 1.680.250 |
| 06/01/15 | Kunyit | Bejo | January | 2015 | Kuartal 1 | 72 | Rp 12.964 | Rp 933.408 |
| 11/01/15 | Bawang Merah | Painah | January | 2015 | Kuartal 1 | 81 | Rp 28.875 | Rp 2.338.875 |
| 18/01/15 | Bawang Putih | Tarno | January | 2015 | Kuartal 1 | 146 | Rp 20.812 | Rp 3.038.552 |
| 22/01/15 | Kencur | Bejo | January | 2015 | Kuartal 1 | 159 | Rp 9.785 | Rp 1.555.815 |
| 29/01/15 | Jahe | Tarno | January | 2015 | Kuartal 1 | 54 | Rp 24.829 | Rp 1.340.766 |
| 04/02/15 | Cabai | Bejo | February | 2015 | Kuartal 1 | 92 | Rp 32.588 | Rp 2.998.096 |
| 06/02/15 | Kunyit | Bejo | February | 2015 | Kuartal 1 | 154 | Rp 10.397 | Rp 1.601.138 |
| 14/02/15 | Bawang Merah | Painah | February | 2015 | Kuartal 1 | 185 | Rp 28.256 | Rp 5.227.360 |
| 18/02/15 | Bawang Putih | Bejo | February | 2015 | Kuartal 1 | 86 | Rp 20.970 | Rp 1.803.420 |
| 22/02/15 | Kencur | Tarno | February | 2015 | Kuartal 1 | 193 | Rp 8.924 | Rp 1.722.332 |
| 26/02/15 | Jahe | Painah | February | 2015 | Kuartal 1 | 156 | Rp 21.751 | Rp 3.393.156 |
| 03/03/15 | Cabai | Painah | March | 2015 | Kuartal 1 | 162 | Rp 28.034 | Rp 4.541.508 |
| 08/03/15 | Kunyit | Bejo | March | 2015 | Kuartal 1 | 148 | Rp 10.231 | Rp 1.514.188 |
| 13/03/15 | Bawang Merah | Tarno | March | 2015 | Kuartal 1 | 140 | Rp 24.792 | Rp 3.470.880 |
| 19/03/15 | Bawang Putih | Bejo | March | 2015 | Kuartal 1 | 81 | Rp 24.323 | Rp 1.970.163 |
| 22/03/15 | Kencur | Tarno | March | 2015 | Kuartal 1 | 175 | Rp 9.733 | Rp 1.703.275 |
| 26/03/15 | Jahe | Tarno | March | 2015 | Kuartal 1 | 66 | Rp 22.275 | Rp 1.470.150 |
| 01/04/15 | Cabai | Bejo | April | 2015 | Kuartal 2 | 135 | Rp 26.969 | Rp 3.640.815 |
| 10/04/15 | Kunyit | Tarno | April | 2015 | Kuartal 2 | 152 | Rp 10.393 | Rp 1.579.736 |
| 15/04/15 | Bawang Merah | Tarno | April | 2015 | Kuartal 2 | 111 | Rp 29.408 | Rp 3.264.288 |
| 16/04/15 | Bawang Putih | Painah | April | 2015 | Kuartal 2 | 60 | Rp 23.416 | Rp 1.404.960 |
| 22/04/15 | Kencur | Painah | April | 2015 | Kuartal 2 | 188 | Rp 9.636 | Rp 1.811.568 |
| 29/04/15 | Jahe | Painah | April | 2015 | Kuartal 2 | 138 | Rp 22.602 | Rp 3.119.076 |
| 04/05/15 | Cabai | Tarno | May | 2015 | Kuartal 2 | 88 | Rp 28.125 | Rp 2.475.000 |
| 07/05/15 | Kunyit | Tarno | May | 2015 | Kuartal 2 | 97 | Rp 12.279 | Rp 1.191.063 |
| 12/05/15 | Bawang Merah | Tarno | May | 2015 | Kuartal 2 | 92 | Rp 26.915 | Rp 2.476.180 |
| 17/05/15 | Bawang Putih | Tarno | May | 2015 | Kuartal 2 | 181 | Rp 21.755 | Rp 3.937.655 |

| | | | | | | | | |
|----------|--------------|--------|------|------|-----------|-----|-----------|--------------|
| 25/05/15 | Kencur | Tarno | May | 2015 | Kuartal 2 | 95 | Rp 9.812 | Rp 932.140 |
| 28/05/15 | Jahe | Painah | May | 2015 | Kuartal 2 | 59 | Rp 22.466 | Rp 1.325.494 |
| 01/06/15 | Cabai | Tarno | June | 2015 | Kuartal 2 | 141 | Rp 31.927 | Rp 4.501.707 |
| 06/06/15 | Kunyit | Bejo | June | 2015 | Kuartal 2 | 190 | Rp 14.491 | Rp 2.753.290 |
| 15/06/15 | Bawang Merah | Bejo | June | 2015 | Kuartal 2 | 74 | Rp 25.083 | Rp 1.856.142 |
| 18/06/15 | Bawang Putih | Painah | June | 2015 | Kuartal 2 | 166 | Rp 21.626 | Rp 3.589.916 |
| 21/06/15 | Kencur | Bejo | June | 2015 | Kuartal 2 | 62 | Rp 9.973 | Rp 618.326 |
| 26/06/15 | Jahe | Tarno | June | 2015 | Kuartal 2 | 54 | Rp 24.649 | Rp 1.331.046 |
| 05/07/15 | Cabai | Bejo | July | 2015 | Kuartal 3 | 148 | Rp 34.622 | Rp 5.124.056 |
| 07/07/15 | Kunyit | Tarno | July | 2015 | Kuartal 3 | 150 | Rp 13.053 | Rp 1.957.950 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |

Catatan: data_bahan_pokok.csv (dataset)

Lakukanlah pivoting tabel menggunakan kode *script* pemrograman Python pembelian bahan pokok makanan yang dikirimkan oleh supplier pada kuartal tertentu menjadi tabel berikut:

Pivoting Tabel Berdasarkan Jumlah Pembelian Bahan Pokok (kg) per Kuartal

| Row Labels | 2015 | 2016 | 2017 | Grand Total |
|---------------------|-------------|-------------|-------------|-------------|
| Bawang Merah | 1450 | 1278 | 1574 | 4302 |
| Kuartal 1 | 406 | 296 | 408 | 1110 |
| Kuartal 2 | 277 | 248 | 427 | 952 |
| Kuartal 3 | 431 | 403 | 430 | 1264 |
| Kuartal 4 | 336 | 331 | 309 | 976 |
| Bawang Putih | 1419 | 1659 | 1601 | 4679 |
| Kuartal 1 | 313 | 534 | 355 | 1202 |
| Kuartal 2 | 407 | 447 | 443 | 1297 |
| Kuartal 3 | 361 | 332 | 506 | 1199 |
| Kuartal 4 | 338 | 346 | 297 | 981 |
| Cabai | 1543 | 1568 | 1612 | 4723 |
| Kuartal 1 | 309 | 272 | 221 | 802 |
| Kuartal 2 | 364 | 473 | 431 | 1268 |
| Kuartal 3 | 519 | 366 | 453 | 1338 |
| Kuartal 4 | 351 | 457 | 507 | 1315 |
| ... | ... | ... | ... | ... |

Pivoting Tabel Berdasarkan Total Pembelian Bahan Pokok (Rupiah) per Kuartal

| Row Labels | 2015 | 2016 | 2017 | Grand Total |
|---------------------|-----------------|-----------------|-----------------|------------------|
| Bawang Merah | 37823531 | 31731767 | 40941357 | 110496655 |
| Kuartal 1 | 11037115 | 6820091 | 10456137 | 28313343 |
| Kuartal 2 | 7596610 | 6251504 | 11075213 | 24923327 |
| Kuartal 3 | 10800386 | 10587768 | 11697016 | 33085170 |
| Kuartal 4 | 8389420 | 8072404 | 7712991 | 24174815 |
| Bawang Putih | 31078455 | 37670606 | 35263522 | 104012583 |
| Kuartal 1 | 6812135 | 12662576 | 7962251 | 27436962 |
| Kuartal 2 | 8932531 | 10457967 | 9784401 | 29174899 |
| Kuartal 3 | 8082531 | 7110937 | 10945924 | 26139392 |
| Kuartal 4 | 7251258 | 7439126 | 6570946 | 21261330 |
| Cabai | 45057879 | 44441269 | 46168221 | 135667369 |
| Kuartal 1 | 9219854 | 7652138 | 5787840 | 22659832 |
| Kuartal 2 | 10617522 | 13099592 | 12230465 | 35947579 |
| Kuartal 3 | 15774604 | 9718410 | 14312881 | 39805895 |
| Kuartal 4 | 9445899 | 13971129 | 13837035 | 37254063 |
| ... | ... | ... | ... | ... |