Mathematical Statistics Problem Set 2
Descriptive Statistics part 2

1. (continuation of Problem 4 from Set 1). The amounts of monthly income taxes paid by workers of a certain factory amounted to (in EUR):

   126.31 140.09 346.72 86.14 106.00 139.72 43.29 79.02 68.35 152.80 75.39 17.35 54.00 114.07 67.08 95.34 71.80 38.77 156.73 204.03 189.38 59.38 17.69 75.43 38.57 153.46 253.92 27.69 95.20 174.10 30.21 104.75 90.97 139.12 291.13 73.49 12.83 248.51 183.43 48.33 10.34 174.59 58.80 33.75 18.32 93.08 73.35 20.18 79.65 268.71 117.49 74.07 97.64 225.97 106.46 60.99 145.19 72.10 27.61 34.75 201.31 125.72 69.42 84.48 106.17 144.38 145.28 56.83 16.97 109.14 39.41 58.02 179.50 28.02 76.67 17.26 15.32 169.48 104.64 36.34 78.94 213.58 15.32 180.15 184.83 107.79 77.17 174.75 32.49 176.68 52.06 68.78 107.16 32.93 141.07 124.24 155.42 2.48 20.75 14.08

   (a) Calculate the variance and the IQR for this group of workers.

   (b) Calculate the variance and the IQR for grouped interval data (with class intervals of length 30, 50 and 70). Discuss the differences and compare with the values calculated for raw data. Should the Sheppard's correction for the variance be used in this case?

   (c) Compare coefficients of asymmetry for raw data and class interval data.

2. A group of students were interrogated on Mardi Gras to determine the number of cakes they had eaten. The results of the survey are summarized in the following table:

number of cakes eaten	0	1	2	3	4	5	6
number of female students	5	10	7	2	1	0	0
number of male students	2	5	6	8	3	0	1

   Compare the Mardi Gras customs of male and female students (central tendency, variation and asymmetry).

   Visualize the results

   (a) by means of bar charts,

   (b) by means of box plots.

3. A group of students were interrogated to determine the amount of time (in minutes) they spent the previous day using electronic devices. The collected data was:

   for female students:
   342, 6, 109, 71, 193, 196, 238, 171, 116, 241, 153, 46, 46, 75, 149, 148, 99, 99, 124, 167

   for male students:
   299, 49, 51, 238, 128, 295, 66, 203, 58, 275, 98, 179, 145, 199, 11, 211, 34, 111, 221, 266

   Compare the habits of male and female students (central tendency, variation and asymmetry).

   Visualize the results

   (a) by means of histograms,

   (b) by means of box plots.

4. The figures below present boxplots of salaries offered in Ukraine in January/February 2017 (source: Sinichenko, Valentyna, Shmihel, Anton and Zhuk, Ivan. (2019). Explaining Wages in Ukraine: Experience or Education?: Proceedings of the XVIII International Conference on Data Science and Intelligent Analysis of Information, June 4-7, 2018, Kyiv, Ukraine) by job type (p.5), required education level (p.5) and city (p.6). Based on these figures, answer the following questions:

   (a) Are there significant differences in the median levels of wage offers for different education levels? Are there differences in the variability of wage offers for different education levels of candidates? Are there differences in the highest wage offer? What is the share of offers for candidates with secondary education with wages higher than 10000? What is the share of offers for candidates with secondary education with wages higher than 5000? When is a wage for candidates with incomplete higher education considered an outlier?

(b) What is the value of the median wage offered for top management positions? What is the IQR of wages offered for jobs in agriculture? Are there significant differences in the median levels of wage offers for different types of jobs? Are there differences in variability of wage offers for different job types? Are there differences in the asymmetry of wage offers for different job types?

(c) Are there significant differences in the wage offers across cities in terms of the median offers, variability or asymmetry? Which city has the lowest first quartile of wage offers?

(d) Where can one find the best offers? Were can one find the worst offers?

(e) What candidates can expect the best offers? What candidates can expect the worst offers?