Last Name:

First Name:

SCIPER:

Practice Problems: Exercise 2 – Microengineering 110

Spring 2025

Prof. Vivek Subramanian

1. The wearing out of some gears was measured in $10^{-4}mm^3$, when using different lubrication oils having different viscosities in mPa-s. The data produced the following table:

Viscosity (mPa-s)	Wear (10^-4 mm^3)					
20.4	58.8	30.8	27.3	29.9	17.7	76.5
30.2	44.5	47.1	48.7	41.6	32.8	18.3
89.4	73.3	57.1	66	93.8	133.2	81.1
252.6	30.6	24.2	16.6	38.9	28.7	23.6

Assuming that outliers are defined by the following rules:

- Lower outlier(s) < $Q1-(1.5\times IQR)$
- Upper outlier(s) > $Q3+(1.5\times IQR)$

a. Calculate the mean and median for each data set (i.e., at each different viscosity)

i. 20.4mPa-s

Mean: 40.2
Median: 30.35

ii. 30.2mPa-s

Mean: 38.8
Median: 43.05

iii. 89.4mPa-s

Mean: 84.1
Medan: 77.2

iv. 252.6mPa-s

Mean: 27.1
Median: 26.45

b. Do any of the data sets have a singular mode?
No. There are no duplicate values.

c. In class, you learned how to calculate Q1 and Q3. One specific issue to consider is what to do when the index falls between two numbers (which is the case in this problem; since there are only 6 data points for each data set, Q1 and Q3 don't fall at an integer location). In the DF.QUANT method shown in class, we interpolate. This is the appropriate method to use for such small data sets.

   Using interpolation, what are the values of Q1, Q3, and IQR for each data set?

   i. 20.4mPa-s

      Q1: 27.95
      Q3: 51.8
      IQR: 23.85

   ii. 30.2mPa-s

      Q1: 35
      Q3:46.45
      IQR: 11.45

   iii. 89.4mPa-s

      Q1: 67.825
      Q3: 90.625
      IQR: 22.8

   iv. 252.6mPa-s

      Q1: 23.75
      Q3: 30.125
      IQR: 6.375

d. Using this method, how many outliers are there in the entire data set (i.e,, the sum of outliers for all viscosities)

   There is 1 outlier. The only outlier is in the 89.4mPa-s data set, where 133.2 is an outlier.

   We can see this from the table below.

Viscosity (mPa-s)	Min	Q1	Q2	Q3	Max	IQR	LO cutoff	UO cutoff	Outliers
20.4	17.7	27.95	30.35	51.8	76.5	23.85	-7.825	87.575	0
30.2	18.3	35	43.05	46.45	48.7	11.45	17.825	63.625	0
89.4	57.1	67.825	77.2	90.625	133.2	22.8	33.625	124.825	1
252.6	16.6	23.75	26.45	30.125	38.9	6.375	14.1875	39.6875	0

e. Another method used to calculate quartiles, which is more appropriate for large data sets, is to round up or down the index when it is a non-integer. This is the method used, for example, at the following website:

https://www.calculatorsoup.com/calculators/statistics/mean-median-mode.php

Repeat the calculations that you did in parts c and d, and determine the number of outliers using this website.

The calculated number of outliers is now 0. This is because the website calculated Q1=66, Q2=77.2, and Q3=93.8. This happened because it rounded 1.5 to be 2 to find the index of Q1, and 4.5 to be 5 when to find the index of Q3.

This method is reasonable when there are large data sets. With small data sets, interpolation is often more reasonable.