Delft University of Technology

Mathematics and sensing technologies for geomatics
GEO1001

Assignment 1

Authors:
Androniki Pavlidou (5267536)

September 21, 2020

1 Part I

1.1 Q.1: Compute mean statistics (mean, variance and standard deviation for each of the sensors variables), what do you observe from the results?

Comparing the values of the mean, variance and standard deviation between the sensors observed that they are in a common range of values. The sensor that shows a difference in the values of the variance and standard deviation is sensor E. As far as the mean is concerned, all sensors have quite close values, which is an expected result taking into account their (almost) same sample, content and the study period (summer months). Since there is a great amount of data it is difficult to draw conclusions for each variable. Here is just an example for Wind Direction:

Sensors	Mean	Variance	Standard deviation
A	209.406	10104.857	100.522
B	183.412	9973.188	99.865
C	183.588	7700.249	87.751
D	198.326	8130.602	90.169
E	223.956	9304.524	96.459

From the table above is noticed that the statistical indexes vary significantly compared with mean with whom both variance and standard deviation are related. In this example, variance's values are quite high in comparison with the mean. High variance shows higher spread /variability in the data. That can be also observed from the values of standard deviation that approximate mean values (high). In addition, a significant variable is the altitude, which has a negative value, indicating that the sensors are located below sea level.

All the above data belong to [2].

```
 === VALUES FOR SENSOR: 1 ====================================================================================
True Direction->          mean: 209.40630048465266       variance: 10104.857537040565    standard deviation: 100.52292045618535
Wind Speed->              mean: 1.290306946688207        variance: 1.2506491788047323    standard deviation: 1.1183242726529423
Crosswind Speed->         mean: 0.9649434571890144       variance: 0.9262185347673694    standard deviation: 0.962402480653166 9
Headwind Speed->          mean: 0.16352988691437803      variance: 1.034522111788517     standard deviation: 1.0171146011087036
Temperature->             mean: 17.96910339256866        variance: 15.857862039390751    standard deviation: 3.982193119298805 6
Globe Temperature->       mean: 21.544588045234246       variance: 68.1638115831204      standard deviation: 8.256137812750003
Wind chill->              mean: 17.838206785137317       variance: 16.257877882926497    standard deviation: 4.032105886869354
Relative humidity->       mean: 78.18477382875606        variance: 375.8581970813183     standard deviation: 19.387062621277064
Heat stress index->       mean: 17.899596122778675       variance: 14.990791436236988    standard deviation: 3.8717943432260173
Dew point->               mean: 13.553877221324719       variance: 9.719544740983556     standard deviation: 3.1176184405702307
Psychro Wet Bulb Temp->   mean: 15.270718901453955       variance: 6.9412225849577585    standard deviation: 2.6346200076970794
Station pressure->        mean: 1016.168255250403 7      variance: 38.45572894292478     standard deviation: 6.20126833340767 3
Barometric pressure->     mean: 1016.1284329563813       variance: 38.4524145072175      standard deviation: 6.201001089115974
Altitude->                mean: -25.98707592891761       variance: 2662.5652610782413    standard deviation: 51.60005097941514
Density Altitude->        mean: 137.31663974151857       variance: 26499.337542182006    standard deviation: 162.7861712252672
NA Wet Bulb Temperature->mean: 15.981542810985461        variance: 10.008064017149449    standard deviation: 3.1635524362889025
WBGT->                    mean: 17.25432148626817        variance: 16.128741421686968    standard deviation: 4.016060435512266
TWL->                     mean: 301.39293214862676       variance: 814.4374985107435     standard deviation: 28.538351362872092
Direction , Mag->         mean: 208.90508885298868       variance: 10101.595596074496    standard deviation: 100.50669428488082
```

Figure 1: Sensor A

Figure 2: Statistical indexes

```
 === VALUES FOR SENSOR: 2 ==========================================================================================
True Direction->          mean: 183.41235864297255        variance: 9973.18819944488        standard deviation: 99.86585101747684
Wind Speed->              mean: 1.242124394184168         variance: 1.300975939291838       standard deviation: 1.1406033224972818
Crosswind Speed->         mean: 0.8356219709208401        variance: 0.8782302674332721      standard deviation: 0.9371394066163647
Headwind Speed->          mean: -0.12980613893376414      variance: 1.25621175563275        standard deviation: 1.1208085276409838
Temperature->             mean: 18.065428109854604        variance: 16.622350826415005      standard deviation: 4.077051732123962
Globe Temperature->       mean: 21.799434571890146        variance: 66.02264103731852       standard deviation: 8.125431744671696
Wind chill->              mean: 17.945920840064623        variance: 17.028945395995418      standard deviation: 4.126614277588277
Relative humidity->       mean: 77.87831179321486         variance: 408.4579746943844       standard deviation: 20.21034326018201
Heat stress index->       mean: 18.00428109854604         variance: 15.432921898366484      standard deviation: 3.928475772913266
Dew point->               mean: 13.530856219709205        variance: 9.632626245886195       standard deviation: 3.103647248945375
Psychro Wet Bulb Temp->   mean: 15.295516962843294        variance: 6.767528367644411       standard deviation: 2.6014473601525
Station pressure->        mean: 1016.6570274636512        variance: 36.82705481507772       standard deviation: 6.068529872636183
Barometric pressure->     mean: 1016.6164781906298        variance: 36.81399341269066       standard deviation: 6.067453618503454
Altitude->                mean: -30.05815831987076        variance: 2544.679977868311       standard deviation: 50.444821120391644
Density Altitude->        mean: 135.58077544426493        variance: 26852.460761272676      standard deviation: 163.86720465447831
NA Wet Bulb Temperature->mean: 15.996809369951535         variance: 9.805292727795887       standard deviation: 3.131340404331009
WBGT->                    mean: 17.321970920840062        variance: 15.82895992807201       standard deviation: 3.978562545451813
TWL->                     mean: 299.45169628432956        variance: 789.7501304021025       standard deviation: 28.102493312909132
Direction , Mag->         mean: 183.2172859450727         variance: 9971.418053377041       standard deviation: 99.85698800473125
```

Figure 3: Sensor B

Figure 4: Statistical indexes

```
 === VALUES FOR SENSOR: 3 ==========================================================================================
True Direction->          mean: 183.58892481810832        variance: 7700.24936804366        standard deviation: 87.75106476871754
Wind Speed->              mean: 1.3714632174616006        variance: 1.4303416746777642      standard deviation: 1.1959689271372247
Crosswind Speed->         mean: 0.9632983023443816        variance: 1.0421533895929145      standard deviation: 1.0208591428757028
Headwind Speed->          mean: -0.2628940986257074       variance: 1.2712181399570899      standard deviation: 1.1274830996325798
Temperature->             mean: 17.91313662085691         variance: 16.09802872264436       standard deviation: 4.012234878798145
Globe Temperature->       mean: 21.587388843977365        variance: 67.91384257555865       standard deviation: 8.240985534240346
Wind chill->              mean: 17.77299919159256         variance: 16.53443667987          standard deviation: 4.0662558551903745
Relative humidity->       mean: 77.96285367825384         variance: 374.47121918069183      standard deviation: 19.35125885260935
Heat stress index->       mean: 17.82825383993533         variance: 15.350046506300938      standard deviation: 3.917913539921592
Dew point->               mean: 13.458124494745352        variance: 10.080073442868077      standard deviation: 3.174913139420995
Psychro Wet Bulb Temp->   mean: 15.196645109135003        variance: 7.236387289573896       standard deviation: 2.690053399019041
Station pressure->        mean: 1016.689329021827         variance: 37.67625638083111       standard deviation: 6.13809875945566
Barometric pressure->     mean: 1016.6518997574777        variance: 37.660394531584316      standard deviation: 6.136806541808558
Altitude->                mean: -30.338722716248988       variance: 2607.4802547953855      standard deviation: 51.063492387373834
Density Altitude->        mean: 129.62287793047696        variance: 26975.694884846056      standard deviation: 164.2427924897956
NA Wet Bulb Temperature->mean: 15.934236054971707         variance: 10.476042928918309      standard deviation: 3.2366715818751692
WBGT->                    mean: 17.22502021018593         variance: 16.540057093366812      standard deviation: 4.066946900731163
TWL->                     mean: 301.8997574777688         variance: 766.2236781950229       standard deviation: 27.68074562209304
Direction , Mag->         mean: 183.08367016976555        variance: 7701.505933821689       standard deviation: 87.75822430873183
```

Figure 5: Sensor C

Figure 6: Statistical indexes

```
 === VALUES FOR SENSOR: 4 ==========================================================================================
True Direction->          mean: 198.32659660468877        variance: 8130.602307980361       standard deviation: 90.16985254496295
Wind Speed->              mean: 1.5816491511721908        variance: 1.739113529289902       standard deviation: 1.3187545371637217
Crosswind Speed->         mean: 1.2105092966855295        variance: 1.450916232128608       standard deviation: 1.2045398424828495
Headwind Speed->          mean: -0.3005658852061439       variance: 1.2320045302185576      standard deviation: 1.109956994760859
Temperature->             mean: 17.99636216653193         variance: 16.099081349837828      standard deviation: 4.012366053818847
Globe Temperature->       mean: 21.359296685529507        variance: 61.17751462256783       standard deviation: 7.821605629445135
Wind chill->              mean: 17.835367825383994        variance: 16.55015978790578       standard deviation: 4.068188760112513
Relative humidity->       mean: 77.94203718674213         variance: 389.6984592290132       standard deviation: 19.740781626597595
Heat stress index->       mean: 17.92162489894907         variance: 15.11153317215288       standard deviation: 3.887355550004769
Dew point->               mean: 13.50860953920776         variance: 10.067811890385965      standard deviation: 3.172981545862813
Psychro Wet Bulb Temp->   mean: 15.26018593371059         variance: 7.041555503019602       standard deviation: 2.6535929422237317
Station pressure->        mean: 1016.7280113177042        variance: 34.973641396799955      standard deviation: 5.913851654953813
Barometric pressure->     mean: 1016.6888843977364        variance: 34.938198997953734      standard deviation: 5.910854337399437
Altitude->                mean: -30.653193209377527       variance: 2418.745529415378       standard deviation: 49.18074348172644
Density Altitude->        mean: 132.41107518189168        variance: 26505.40781590138       standard deviation: 162.80481508819506
NA Wet Bulb Temperature->mean: 15.915642683912694         variance: 9.983397181945264       standard deviation: 3.1596514336149903
WBGT->                    mean: 17.1767987065481          variance: 15.500916833369384      standard deviation: 3.9371203732384643
TWL->                     mean: 305.254567502021          variance: 615.7608138186043       standard deviation: 24.814528281202612
Direction , Mag->         mean: 197.8261924009701         variance: 8132.027187846571       standard deviation: 90.17775328675344
```

Figure 7: Sensor D

Figure 8: Statistical indexes

```
 === VALUES FOR SENSOR: 5 ==================================================================================================
True Direction->          mean: 223.95636363636365        variance: 9304.524156473828     standard deviation: 96.459961416506
Wind Speed->              mean: 0.5962424242424242        variance: 0.5110202240587696    standard deviation: 0.7148567856982051
Crosswind Speed->         mean: 0.4385050505050505        variance: 0.3158143307825732    standard deviation: 0.5619736032791693
Headwind Speed->          mean: 0.19494949494949496       variance: 0.3189441893684318    standard deviation: 0.5647514403420604
Temperature->             mean: 18.353939393939395        variance: 19.035438016528925    standard deviation: 4.362962069114161
Globe Temperature->       mean: 21.176161616161615        variance: 63.18996102438527     standard deviation: 7.949211346063537
Wind chill->              mean: 18.2940202020202          variance: 19.129329898581776    standard deviation: 4.373708940771182
Relative humidity->       mean: 76.7930505050505          variance: 406.3302224115906     standard deviation: 20.157634345616813
Heat stress index->       mean: 18.286424242424246        variance: 18.46777529476584     standard deviation: 4.297414954919509
Dew point->               mean: 13.558787878787879        variance: 9.418778328741965     standard deviation: 3.0690028231889857
Psychro Wet Bulb Temp->   mean: 15.406666666666666        variance: 6.994618181818182     standard deviation: 2.6447340474645427
Station pressure->        mean: 1016.1661010101009        variance: 38.92418015141312     standard deviation: 6.23892459895238
Barometric pressure->     mean: 1016.127797979798         variance: 38.91944545413724     standard deviation: 6.238545139224148
Altitude->                mean: -25.96121212121212        variance: 2691.26556620753      standard deviation: 51.877409015943826
Density Altitude->        mean: 150.84     variance: 29702.92147070707      standard deviation: 172.3453552339229
NA Wet Bulb Temperature->mean: 15.93688888888889         variance: 9.428372543209877     standard deviation: 3.0705655086986625
WBGT->                    mean: 17.185535353535354        variance: 15.483612996224876    standard deviation: 3.934922235092439
TWL->                     mean: 284.11531313131314        variance: 1289.3922059120498    standard deviation: 35.908107801888555
Direction , Mag->         mean: 223.89656565656566        variance: 9264.263240730537     standard deviation: 96.2510428033407
```

Figure 9: Sensor E

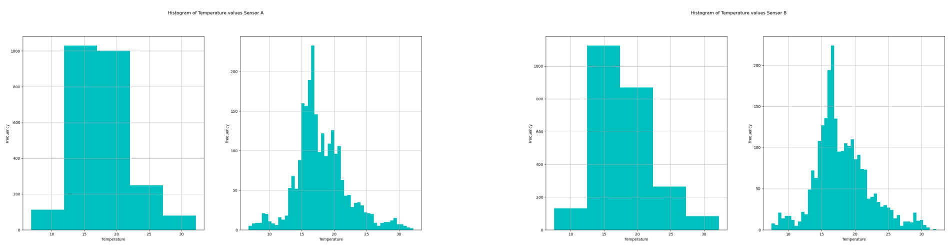
Figure 10: Statistical indexes

1.2 Q.2: Create 1 plot that contains histograms for the 5 sensors Temperature values. Compare histograms with 5 and 50 bins, why is the number of bins important?

Once the required histograms created, the results regarding the importance of bins are the following:

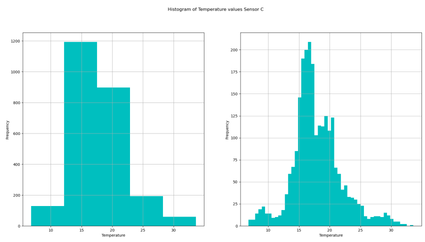
- The wider the range we used, the fewer columns we had
- When bins = 5 we hide important details about distribution while bins = 50 cause a lot of noise and important information about the distribution as well.
- Using higher number of bins gives the viewer the ability to distinguish global or local maximums that helps to understand the data's' fluctuation.
- It is important to use a logical number of bins (depending on the specifications of each project) so on to be easier for the viewer to interpret the data.
- Using too narrow bins results to lots of spikes just by coincidence.

In general, if we have a small amount of data it is preferable to use wider bins to eliminate noise, otherwise, narrower because the histogram will not be that noisy.

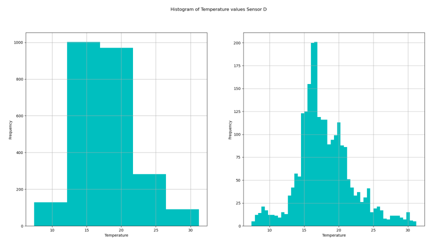


(a) Sensor A (b) Sensor B

Figure 11: Temperature histograms: bins = 5, bins = 50



(a) Sensor C.



(b) Sensor D

Figure 12: Temperature histograms: bins = 5, bins = 50

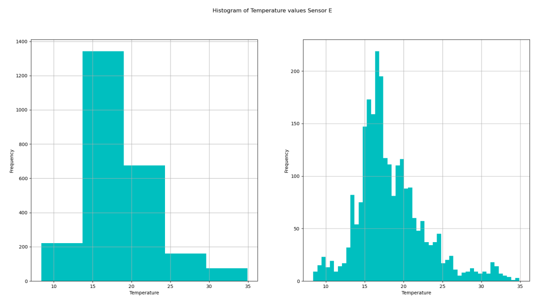


Figure 13: Sensor E

Figure 14: Tempreture histogram with bins=5 , bins=50

1.3 Q.3: Create 1 plot where frequency poligons for the 5 sensors Temperature values overlap in different colors with a legend.

In question A1:Q3, a frequency polygons plot created for all the sensors, for the “Temperature” variable. The results obtained are as follows:

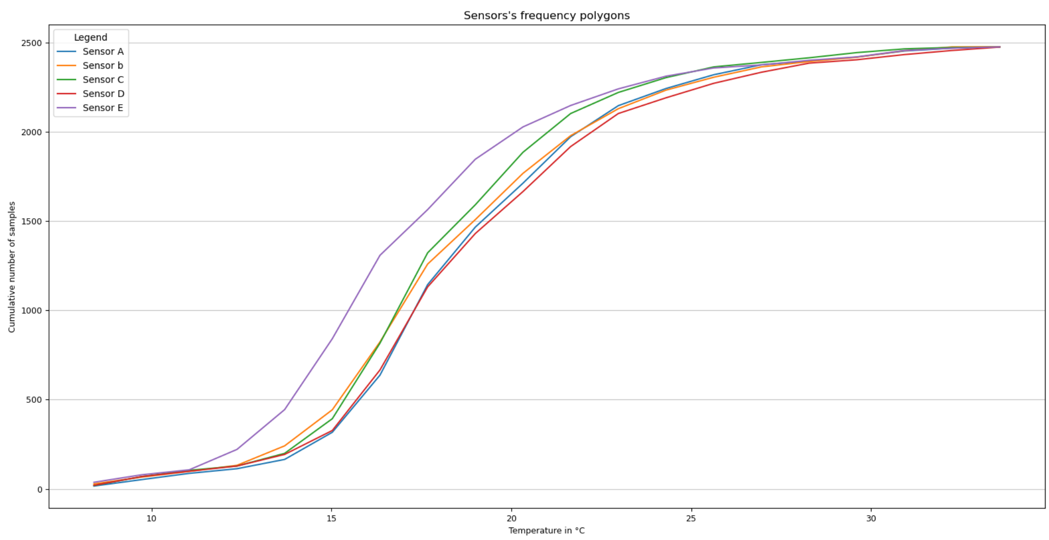


Figure 15: Frequenncy polygon plot for Temperature

Through the frequency's polygon plot we can understand the shape of the distribution and they are helpful for comparing the different sets of data. The given plot illustrates information about the frequency of each sensor for the variable of Temperature. It is clear that the majority of sensors follow a same temperature pattern expect from sensor E that seems to have its highest percentage of (temperature) values between 15°C and 24 °C.

1.4 Q.4: Generate 3 plots that include the 5 sensors boxplot for: Wind Speed, Wind Direction and Temperature.

Another way to represent visually our data is boxplots. In this case, apart from comparing distributions we receive also important information for the outliers. Through boxplots and its' elements (minimum, maximum, sample median, and first and third quartiles) we can understand the behavior of a distribution. For example through Winds' Direction boxplot and its y-label values we can estimate winds' cardinal direction an important information that will helps us, as we will see below, to discern the possible positions of the sensors. Moreover, the former boxplot depicts information about the skewness of the distribution. Comparing the mean with the median helps us to clarify if the distribution has a positive or negative skew (mean>median –> positive skew).

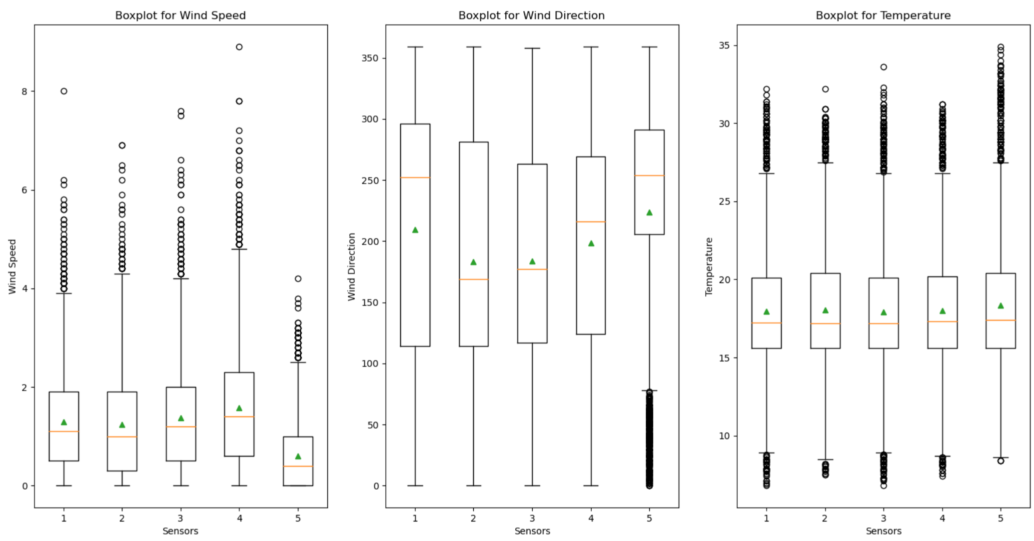


Figure 16: Wind Speed, Wind Direction and Temperature boxplots

2 Part II

2.1 Q.1: Plot PMF, PDF and CDF for the 5 sensors Temperature values in independent plots (or subplots). Describe the behaviour of the distributions, are they all similar? what about their tails?

Comparing the PMF, PDF and CDF temperature graphs for each sensor we see that they show significant similarities and follow similar patterns. Probability Mass Distribution, in general, gives the probability that a discrete random variable is exactly equal to some value. Probability Density Function of a continuous random variable, is a function whose value at any given sample (or point) in the sample space (the set of possible values taken by the random variable) can be interpreted as providing a relative likelihood that the value of the random variable would equal that sample. Regarding PMF and PDF we see that they are quite similar. They are higher in the middle compared to its two tails. Their tail in the positive direction extends further than the tail in the negative one (right-skewed). Moreover they both range from almost 4°C to 35°C and reach a top of 0.024 and 0.016, PMF and PDF respectively. As far as CDF is concerned, is the probability that the variable takes a value less than or equal to x. In this plot is observed that there are no significant changes in CDFs form whilst all sensors appear to follow a similar pattern of distribution on their temperatures.

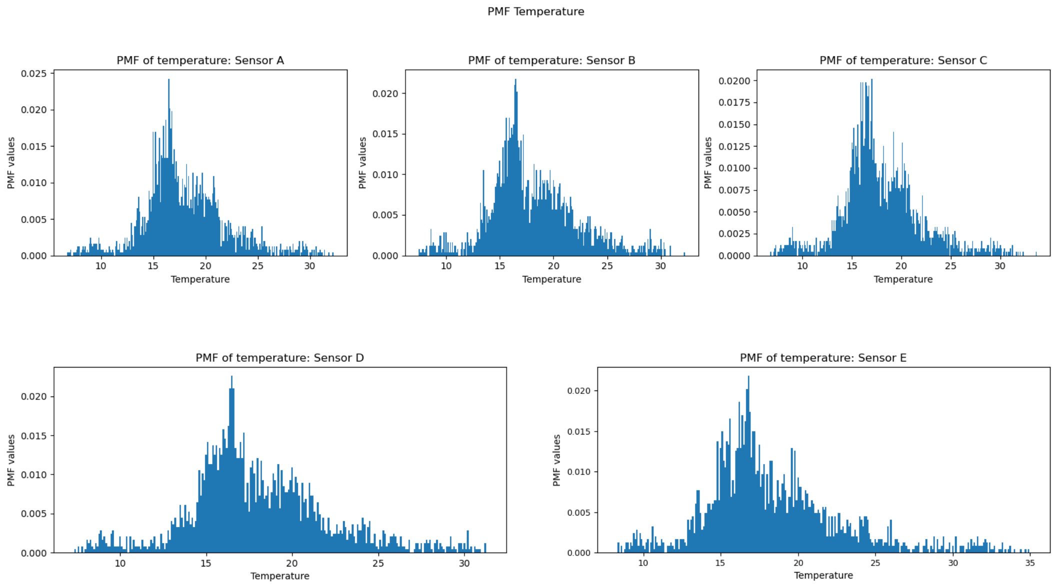


Figure 17: Probability Mass Function for Temperature

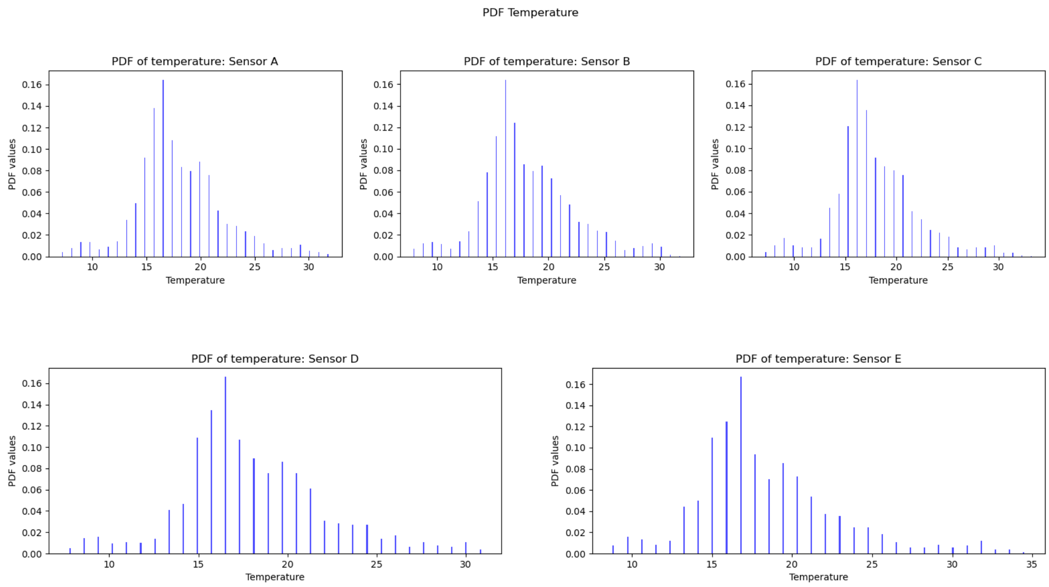


Figure 18: Probability Density Function

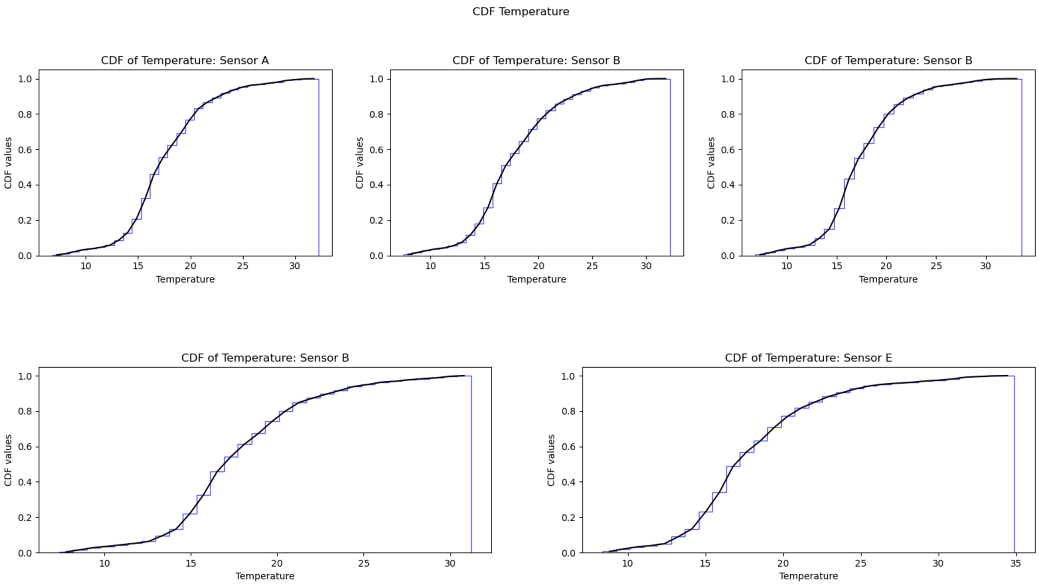


Figure 19: Cumulative Density Function

2.2 Q.2:For the Wind Speed values, plot the pdf and the kernel density estimation. Comment the differences.

Analyzing Kernel's Density Estimation plot we noticed a smoother presentation of the data, a fact that results from KDEs way of working (plotting out the data and beginning to create a curve of the distribution). Since KDE is an algorithm that estimates a PDF (based on a sample) the main difference between them is that the former one gives a general visualization of the data and this may cause the lost of some extreme values. The majority of sensors (A, B, C, D) shows a range of wind speed between 0 and 8 [m/s], apart from sensor E, who did not surpass the speed of 3 [m/s]. Regarding the wind speed values we can estimate that sensor E is located at a point which is not so affected from the wind.

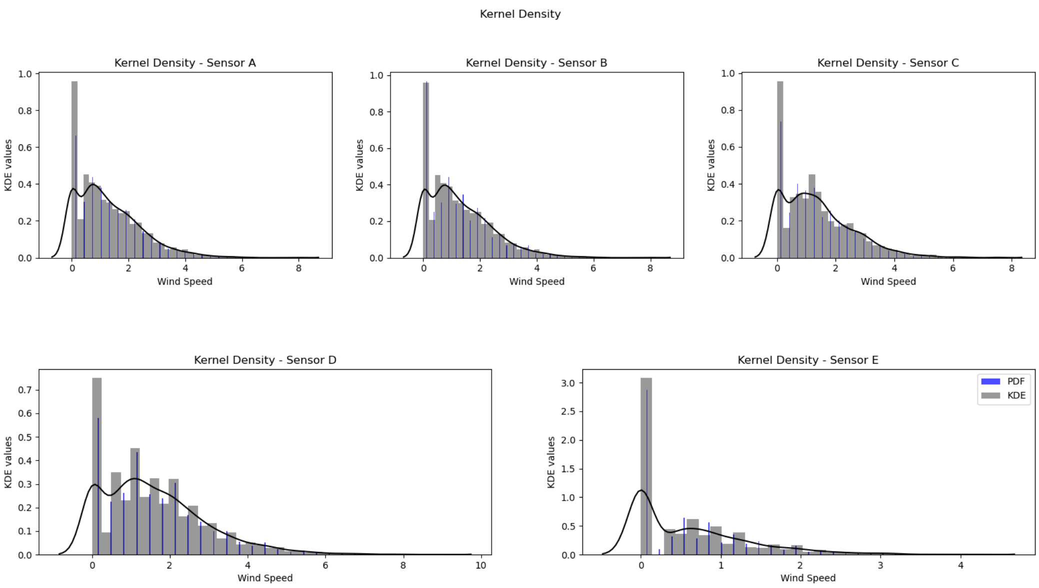


Figure 20: Kernel Density for Wind Speed

Part III

2.3 Q.1:Pearson's - Spearman's correlations

Question A1:Compute the correlations between all the sensors for the variables: Temperature, Wet Bulb Globe Temperature (WBGT), Crosswind Speed. Perform correlation between sensors with the same variable, not between two different variables; for example, correlate Temperature time series between sensor A and B. Use Pearson's and Spearmann's rank coefficients. Make a scatter plot with both coefficients with the 3 variables. Question A1.1: What can you say about the sensors' correlations? Question A1.2: If we told you that that the sensors are located as follows, hypothesize which location would you assign to each sensor and reason your hypothesis using the correlations.

In order to quantify the strength of the relationship between the requested variables (Temperature, Wet Bulb Globe, Crosswind) for all the sensors, Pearson's and Spearman's rank coefficients computed. With Pearson's correlation each value transformed to a standard score, which is the number of standard deviations from the mean since with Spearman's correlation each value transformed to its rank, which is its index in the sorted list of values. Comparing the outcomes of both rank coefficients for all sensors we observed that the combinations of the sensors A, B, C, D with E show less correlation, with AE and BE having the lowest one, concerning the variables of Temperature and Crosswind respectively. For the tested variables sensors A-C, B-D have the strongest positive connection (about 0.98) which is an indication useful for finding the possible location of the sensors. On Cross Wind Speed scatter plot Spearman's rank coefficients fluctuated between 0.530 – 0.600 whilst Pearson's between 0.400-0.560. These values differ significantly compared to the correlation values shown in the Temperature and Wet Bulb Globe scatter plots. However, both values represent also strong correlations and in this case is observed that sensors B-E have the lowest correlations (comparing with the other values).

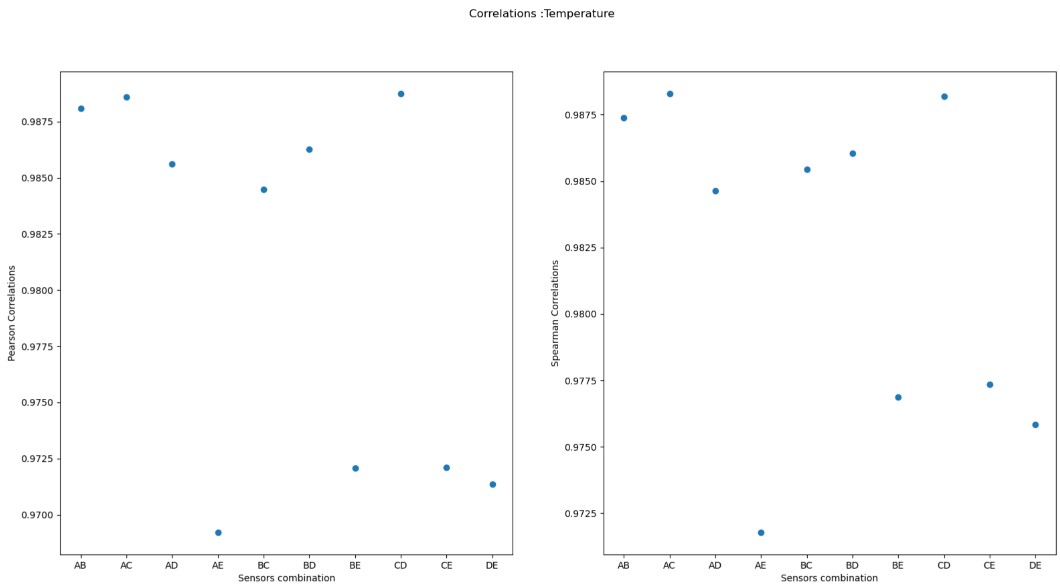


Figure 21: Pearson's and Spearman's rank coefficients: Temperature

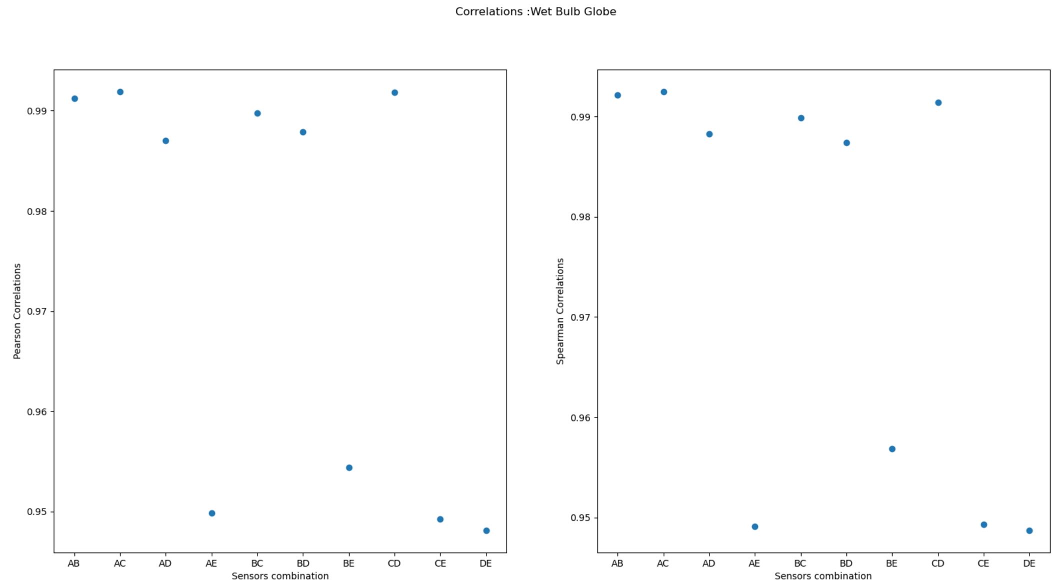


Figure 22: Pearson's and Spearman's rank coefficients: Wet Bulb Globe

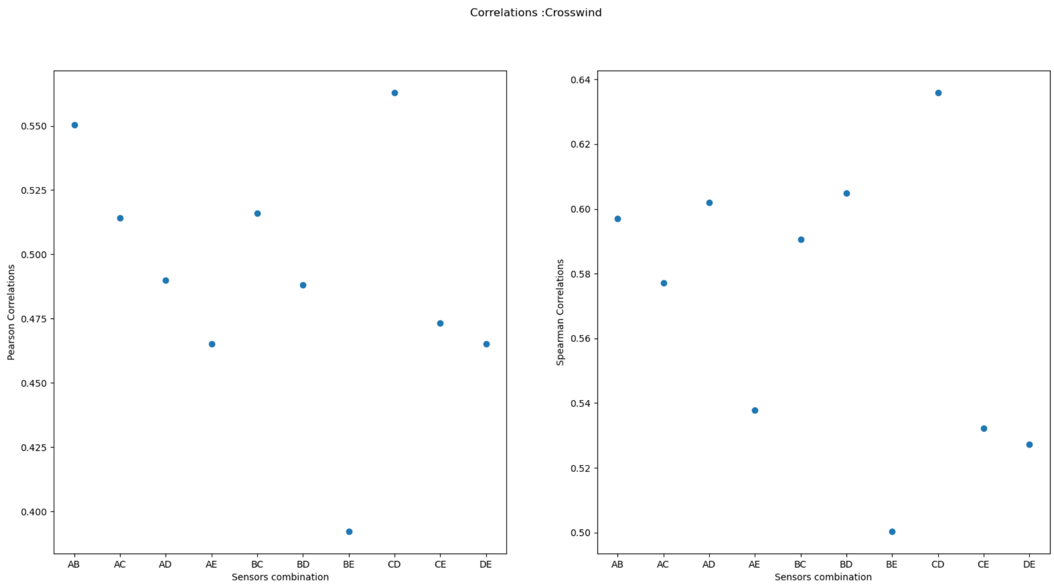


Figure 23: Pearson's and Spearman's rank coefficients: Crosswind Speed

Sensors	Pearson's Rank Coefficients Temperature	Spearman's Rank Coefficients Temperature	Pearson's Rank Coefficients Wet Bulb Globe	Spearman's Rank Coefficients Wet Bulb Globe	Pearson's Rank Coefficients Crosswind Speed	Spearman's Rank Coefficients Crosswind Speed
AB	0.9880961160961126	0.9873789546525072	0.9912595533881619	0.9921324359540058	0.5503525849570334	0.5969825624049757
AC	0.9886087185252328	0.9882920066209426	0.9918958502071857	0.9924720182971508	0.5140508798931707	0.5772288910798762
AD	0.985613462024903	0.9846272388693882	0.9870139489166734	0.9882919234478525	0.4898950130186947	0.6018890586328168
AE	0.9692047916162694	0.9717698000821421	0.9498286924654165	0.9491275351688659	0.4651246851197132	0.5378446650454964
BC	0.9844851698356611	0.9854401094930247	0.989729693523277	0.9898635757569907	0.5161024168073268	0.5906839137964633
BD	0.9862654029844026	0.9860487230587479	0.9878642090483689	0.9873748114350143	0.48802933817375793	0.604818772469813
BE	0.9720897382360566	0.976859613455202	0.9544089298174472	0.9569004735371843	0.39214871017571246	0.500281381925235
CD	0.9887428724207228	0.9881855891390963	0.9918205586342298	0.9914219338897717	0.5628881993613148	0.6359061682587346
CE	0.9720972146615451	0.9773424118180742	0.9492695317424182	0.9493455874960454	0.47323322832012416	0.5322320929791761
DE	1, 0.9713657061948087	0.9758482550943606	0.9480902122116064	0.9487020195244296	0.46519207768485027	0.5273253265612854

Considering the three (3) scatter plots and taking into account the results from all the above diagrams (histograms, boxplot) we can estimate sensors's possible position. It is noted that since the differences between the coefficients are almost fractional there is more than one possible positions for sensors A, B, C and D, since for sensor E all the data we have leads to the initial assumption that is located in a remote place.

The following Figure (24) shows the possible sensors's position.

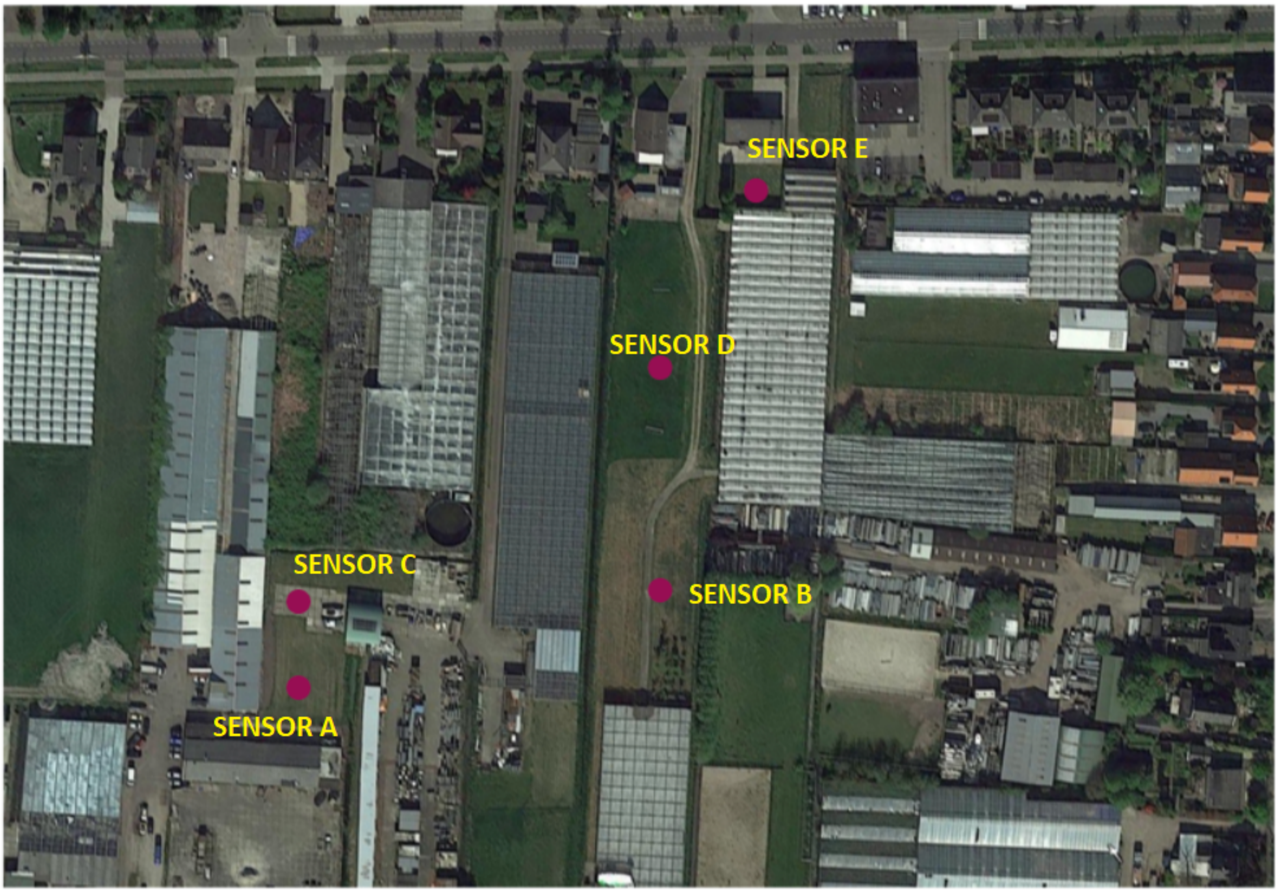


Figure 24: Sensors possible location

3 Part IV

3.1 Q.1: Plot the CDF for all the sensors and for variables Temperature and Wind Speed, then compute the 95/100 confidence intervals for variables Temperature and Wind Speed for all the sensors and save them in a table (txt or csv form).

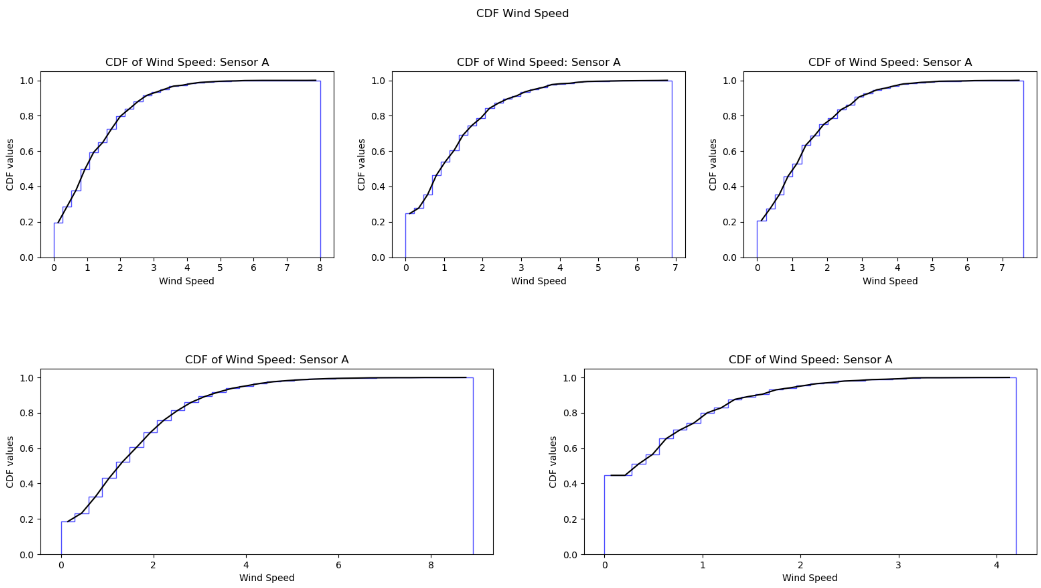


Figure 25: Cumulative Density Function for Wind Speed

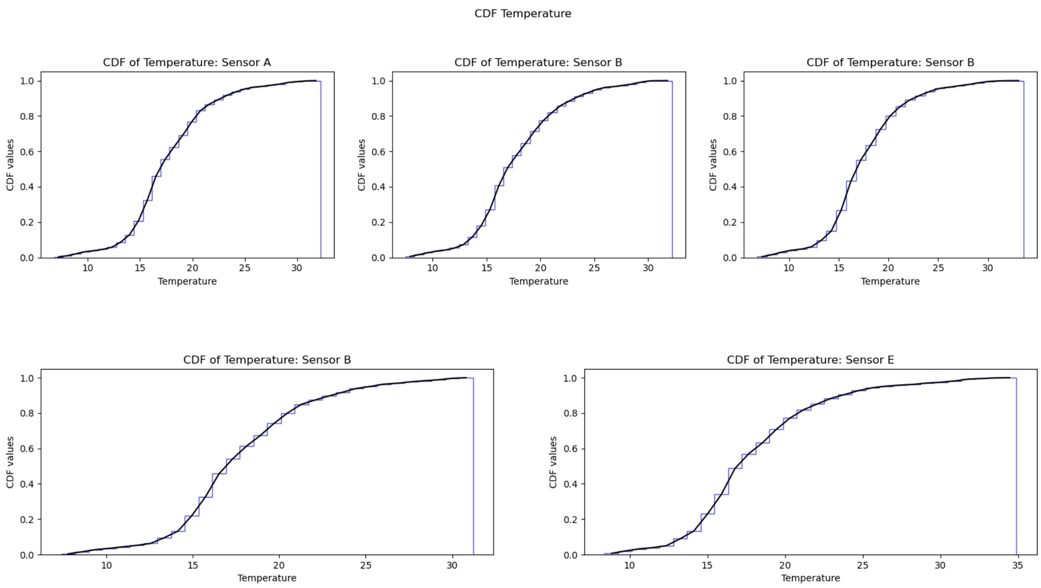


Figure 26: Cumulative Density Function for Temperature

Sensors	TEMPERATURE		WIND SPEED	
	σ1	σ2	σ1	σ2
A	17.812141132673453	18.12606565246385	1.2462270389909704	1.334386854385442
B	17.90472689963894	18.226129320070267	1.1971663346979255	1.2870824536704117
C	17.754926235060218	18.071347006653546	1.324303788594895	1.4186226463283098
D	17.83814660824381	18.15457772482005	1.5296480419653766	1.6336502603790068
E	18.181933946027776	18.525944841851015	0.5680599051948442	0.6244249432900045

Figure 27: 95/100 Confidence intervals for Temperature and Wind Speed

3.2 Q.2: Test the hypothesis: the time series for Temperature and Wind Speed are the same for sensors: 1) E, D; 2) D, C; 3) C, B; 4) B, A. What could you conclude from the p-values?

In order to draw conclusions regarding p-values for the variables of Temperature and Wind Speed a null hypothesis should specified. In our case the former hypothesis is "the time series for each variable is the same for all sensors". That means:

$$x_1 - x_2 = 0 \tag{1}$$

x_1: Temperature values for sensor E
x_2: Temperature values for sensor D

Since the hypothesis test determines that the times series are the same ($x_1 = x_2$) it will be used the two-tailed probability. For the test was specified the α level (significance level) to be equal to $\alpha = 0.05$ (95/100 confidence intervals). The next step is the computation of the probability value. The results are the following:

Sensors	p - value Wind Speed		Significance level	Result
E , D	0.002711172129731209	<	0.05	Null Hypothesis Rejected
D , C	0.4657972008220813	>	0.05	~~Alternative Hypothesis~~
C , B	0.1854863671761938	>	0.05	~~Alternative Hypothesis~~
B , A	0.4004754260262924	>	0.05	~~Alternative Hypothesis~~

(a) Probability values for Wind Speed measurements

Sensors	p - value Temperature		Significance level	Result
E , D	3.3729639501474365e-212	<	0.05	Null Hypothesis Rejected
D , C	4.610149126224334e-09	<	0.05	Null Hypothesis Rejected
C , B	0.00010045473692816457	<	0.05	Null Hypothesis Rejected
B , A	0.13351922750703515	>	0.05	~~Alternative Hypothesis~~

(b) Probability values for Temperature measurements

Figure 28: P - values

In order to accept or reject the null hypothesis the above results compared with the α level. If the former value is lower than the latter one the null hypothesis rejected (stastical significant). The lower the probability value, the more confidence you can have that the null hypothesis is false and so there is stronger evidence in favor of the alternative hypothesis [1].The alternative hypothesis states whether the population parameter differs from the value of the population parameter stated in the conjecture. On the other hand, failure to reject the null hypothesis means that you do not have sufficiently strong data to reject it (we cannot conclude that a significant difference exist or we accept directly the opposite).

4 Bonus question: Your "employer" wants to estimate the day of maximum and minimum potential energy consumption due to air conditioning usage. To hypothesize regarding those days, you are asked to identify the hottest and coolest day of the measurement time series provided. How would you do that? Reason and program the python rutine that would allow you to identify those days.

In order to estimate the hottest and coolest day of the given measurement time series there should be followed the next steps:

- Data control: focus on date-time-measurements for the Temperature variable
- Computation of mean temperatures/day: at this point, the averages of the daily temperatures calculated in order then to identify their maximum and minimum values, for the study period.

Since the employer does not determine a certain date / hour frame of calculations (ex. specific hours of the day, morning – night etc.), determining the average temperature per day and then the minimum and maximum values, seems the most logical solution.

Sensors	Hotter Temperature (degrees)	Coolest Temperature degrees	Hotter Day	Coolest Day
A	25.183	14.155	2020-06-26	2020-06-10
B	24.929	14.327	2020-06-26	2020-06-10
C	24.872	14.266	2020-06-26	2020-06-10
D	24.875	14.370	2020-06-26	2020-06-10
E	25.911	14.490	2020-06-25	2020-07-08

Figure 29: Hypothesize the day of maximum and minimum potential energy consumption due to air conditioning usage

From the table above is noticed that the sensors A, B, C and D have a maximum temperature at almost 25°C in 2020-06-26 whilst sensor E has its highest value of temperature one day before (2020-06-25). On the other hand, the lowest temperature that sensors had at the study period was 14°C. Sensors A, B, C and D had that temperature in 2020-06-10 but as far as E sensor is concerned that temperature were reached in 2020-07-08. Regarding these results, we can estimate that the day with the maximum potential energy consumption due to air conditioning usage is 2020-06-25, since that day was the hottest one whilst the day of minimum energy consumption is 2020-07-08 (there is not a high need of air conditioning usage). In general when we have the greatest difference between indoor and outdoor temperatures then we have the highest energy consumption due to air conditioning usage. The exact day can be determined when both of the above parameters are known.

References

[1] J. Bruin. newtest: command to compute new test @ONLINE, February 2011.

[2] Daniela Maiullari and Clara Garcia Sanchez. Measured Climate Data in Rijsenhout. 8 2020.