# Lab 3 | Statistics Foundations

In this lab we are going to put into practice what we learned about the foundations of statistics. You won't need to use Python, just your brain and a little bit of \*Markdown\*. Hovewer, you will need SQL for the Part 3.

## Your task

Today you'll need to complete the challenges described below.

## Deliverables

You need to submit a markdown file with the solution to the following challenges. You can create a new \*.md\* file or directly edit the \*README.md\* to include your solutions.

This file should also include the SQL script for the Part 3.

## Challenges

**## Part 1**

**### Challenge 1**

Find a and b, if median=42, mean=50

20,34, 34, 12, 56, 110, 49, a, b.

Sum[12, 20, 34, 34, 49, 56, 110] = 315

(315 + a +b) / 9 = 50

So a+b = 450 – 315 = 135

Full set has 9 figures so the median is the figure #5 !

The median before adding in a & b was 34.

Which implies a and b are both > 34

And one of a or b = 42

If a = 42 then b = 135 – 42 = 93

**## Challenge 2**

A car hit an average speed of v1 = 60 kmph in the first half of the journey, and hit an average speed of v2 = 110 kmph in the second half of the journey. Determine the average speed of the car.

Let’s say car traveled 1 hour at 60kmph and another hour at 110kmph.

Total distance traveled was 60 + 110 km in 2 hours

Average speed was then 170 / 2 = 85 kmph

**### Challenge 3**

The lifetimes of 400 light-bulbs were found to the nearest hour. The results were recorded as

follows.

Lifetime (hours) 0–199 200–399 400–599 600–799 800–999 1000–1199 1200–1999

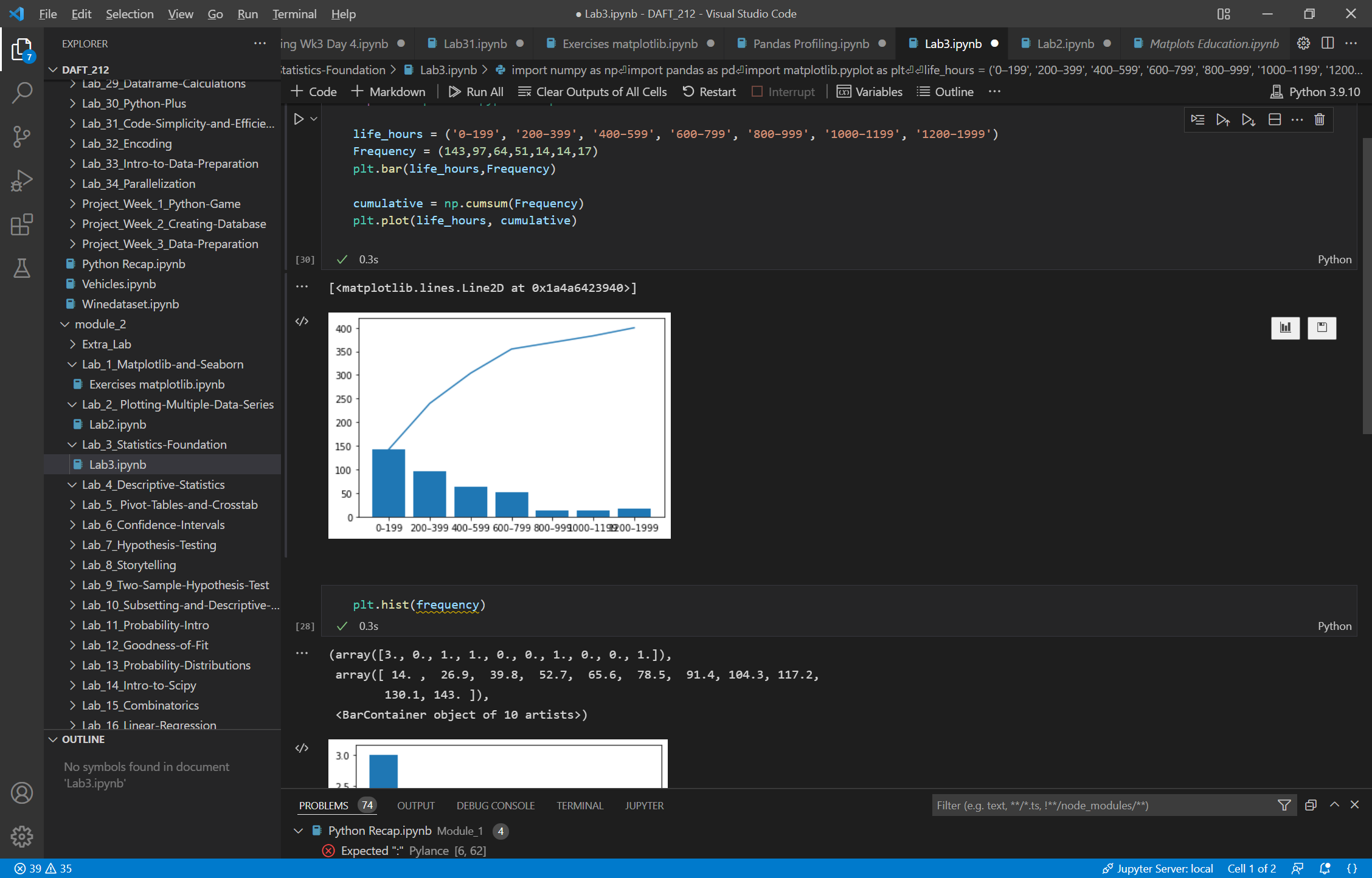
Frequency 143 97 64 51 14 14 17

Construct a histogram and cumulative frequency polygon for this dataset. Estimate the percentage

of bulbs with lifetime less than 480 hours.

Cumulative frequency up to 480 hours = 143+97+ (64)\*80/200 = 266

% of bulbs with lifetime < 480 hrs = 266/400 = 66.5%



**### Challenge 4**

The time between arrival of 60 patients at an intensive care unit were recorded to the nearest hour.

The data are shown below.

Time (hours) 0–19 20–39 40–59 60–79 80–99 100–119 120–139 140–159 160–179

Frequency 16 13 17 4 4 3 1 1 1

Determine the median, mean and standard deviation for this dataset.

Median = 40 + (60/2 – 29) / 17 \*20 = 41.18

Mean = 6.6667

Std Dev = 2.44

|  |  |  |
| --- | --- | --- |
| 16 | 6,66666 | 87,1112356 |
| 13 | 6,66666 | 40,1111956 |
| 17 | 6,66666 | 106,777916 |
| 4 | 6,66666 | 7,11107556 |
| 4 | 6,66666 | 7,11107556 |
| 3 | 6,66666 | 13,4443956 |
| 1 | 6,66666 | 32,1110356 |
| 1 | 6,66666 | 32,1110356 |
| 1 | 6,66666 | 32,1110356 |
|  |  |  |
|  | Sum of squared gap vs mean | 358 |
|  | Var | 5,96666667 |
|  | Std Dev | 2,44267613 |

**## Part 2**

**### Challenge 1**

One player rolls two dices. Describe the measurable space and the random variable for:

\* A. The values that the player obtains.

\* B. The sum of the values obtained.

\* C. The maximum value obtained after rolling both dices.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 2 dice | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 2 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 3 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 4 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 5 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 6 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sum of values | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 6 | 7 | 8 | 9 | 10 | 11 | 12 |

Max value = 12

Describe the following events:

\* Case A: Both values are greater than 5.

Probability = 4 out of 36

\* Case B: The sum of values is even.

Probability = 18 out of 36

\* Case C: The maximum is the value of both rolls.

Probability = 1 out of 36

**### Challenge 2**

One player picks two cards from a poker deck. Describe the measurable space and the random variable for:

\* A. The number of figures he picks.

If figure is one of the 3 figures (J, Q, K), probability of drawing one figure = 12 /52

Second card figure probability = 11/51

P(2 figures) = 12/52 \* 11/51

P(1 figure) = 12/52 \* 41/51

\* B. The sum of card values. Consider that the value of figures is 10 and the value of aces is 15.

\* C. The number of hearts or spades he picks.

First card probability being hearts = P1 = 13/52

Second card probability being hearts = P2 = 12/51

Probability of 2 hearts = P1 \* P2

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2 cards | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | J | Q | K | A |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| J |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Q |  |  |  |  |  |  |  |  |  |  |  |  |  |
| K |  |  |  |  |  |  |  |  |  |  |  |  |  |
| A |  |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sum of cards | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **10** | **10** | **10** | **15** |
| **2** | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 12 | 12 | 12 | 17 |
| **3** | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 13 | 13 | 13 | 18 |
| **4** | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 14 | 14 | 14 | 19 |
| **5** | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 15 | 15 | 15 | 20 |
| **6** | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 16 | 16 | 16 | 21 |
| **7** | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 17 | 17 | 17 | 22 |
| **8** | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 18 | 18 | 18 | 23 |
| **9** | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 19 | 19 | 19 | 24 |
| **10** | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 20 | 20 | 20 | 25 |
| **10** | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 20 | 20 | 20 | 25 |
| **10** | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 20 | 20 | 20 | 25 |
| **10** | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 20 | 20 | 20 | 25 |
| **15** | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 25 | 25 | 25 | 30 |

13 cards of each color

Describe the following events:

\* Case A: The number of figures in the cards the player picked is two.

Probability of 2 figures = 12/52 \* 11/51

\* Case B: The sum of card values is 17.

17 = 10+7 = 7+10 = 15+2 = 2+15 = 9+8 = 8+9

P(17) = (4/52\*4/51 + 4/52\*4/51 +4/52\*4/51)\*2 = 3,6%

\* Case C: The value of both cards is less than 8.

P(case C) = 24/52\*23/51 = 552 / 2652 = 20,8%

### Challenge 3

Two players roll a dice. Describe the measurable space and the random variable for:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Player A (rows) vs Player B (cols) | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 2 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 3 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 4 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 5 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 6 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Points gap | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 0 | 1 | 2 | 3 | 4 | 5 |
| 2 | 1 | 0 | 1 | 2 | 3 | 4 |
| 3 | 2 | 1 | 0 | 1 | 2 | 3 |
| 4 | 3 | 2 | 1 | 0 | 1 | 2 |
| 5 | 4 | 3 | 2 | 1 | 0 | 1 |
| 6 | 5 | 4 | 3 | 2 | 1 | 0 |

\* A. The score of player A.

6 possible outcomes from 1 to 6

Each outcome probability being 1/6

\* B. The greatest score.

Greatest possible score = 6 … with a probability of 10/36

Greatest possible score = 5 … with a probability of 8/36

Greatest possible score = 4 … with a probability of 6/36

Greatest possible score = 3 … with a probability of 4/36

Greatest possible score = 2 … Probability = 2/36

Greatest possible score = 1… Probability = 1/36

\* C. The earnings of player A if the game rules state that:

"The player with the greatest score gets a coin from the other player.".

Probability of A getting a coin (higher score than B) = 15/36

\* D. The earnings of player A if the game rules state that:

"The player with the greatest score gets as many coins as the difference between the score of player A and player B.".

Player A probility to earn coins:

P(1 coin) = 5/36

P(2 coins)= 4/36

P(3 coins)= 3/36

P(4 coins) = 2/36

P(5 coins) = 1/36

Describe the following events:

\* Case A: The score of player A is 2.

P(case A) = 6/36

\* Case B: The greatest score is lower or equal than 2.

P(case B)= 4/36

\* Case C: Considering the case where the winner gets as many coins as the difference between scores (D), describe:

\* Player A wins at least 4 coins. Probability = 3/36

\* Player A loses more than 2 coins. Probability = 6/36

\* Player A neither wins nor loses coins. Proba = 6/36

## Bonus challenges

### Bonus Challenge 1

Three players take balls from a box. Inside that box there are red, blue, green and black balls. The players can take three balls at mosts with the following rules:

\* If the ball is blue, they can take another ball.

\* If the ball is green, they get one point and they can take another ball.

\* If the ball is red, they can’t take another ball.

\* If the ball is black, they lose one point and they can’t take another ball.

Describe the measurable space and the random variable for:

\* A. Player A wins. Do not consider ties as a win.

\* B. Player A and B get the same points.

\* C. All players get 0 points.

### Bonus Challenge 2

Consider the situation of bonus challenge 1 but now with four players. Does anything change in your solutions? What are the changes in each case?

### Bonus Challenge 3

One player takes three balls from a box. Inside the box there are 5 balls: two of them are black and the other three are white.

Describe the measurable space and the random variable for:

\* A. The number of white balls if every time we take a ball we keep it.

\* B. The number of white balls if every time we take a ball we put it back again into the box.

\* C. The number of black balls if every time we take a ball we keep it.

\* D. The number of black balls if every time we take a ball we put it back into the box.

| Attempt | #1 | #2 |

| :---: | :---: | :---: |

| Seconds | 301 | 283 |

## Part 3

### Challenge 1

| #student\_id | #score |

| :---------: | :----: |

| 1 | 91 |

| 2 | 72 |

| 3 | 98 |

| 4 | 62 |

| 5 | 62 |

| 6 | 95 |

| 7 | 83 |

| 8 | 86 |

| 9 | 56 |

|10 | 97 |

|11 | 58 |

|12 | 71 |

|13 | 87 |

|14 | 83 |

|15 | 98 |

Please find MEAN, SUM, STANDARD DEVIATION, VARIANCE, MIN and MAX.

|  |  |
| --- | --- |
| mean | 79,93 |
| mode | 98 |
| sum | 1199 |
| var | 215,9 |
| std dev | 14,69 |
| min | 56 |
| max | 98 |