

## **Lab report 3**



**Fall 2021**

**CSE422L Data Analytics Lab**

Submitted by: **Ayaz Mehmood**

Registration No.: **18PWCSE1652**

Section: **A**

“On my honor, as student of University of Engineering and Technology, I have neither given nor received unauthorized assistance on this academic work.”

Student Signature: \_\_\_\_\_

Submitted to:

**Engr. Mian Ibad Ali Shah**

Last date of Submission:

**November 2021**

**Department of Computer Systems Engineering**  
**University of Engineering and Technology, Peshawar**

## OBJECTIVE:

The basic Objective of this lab is:

- To know about the distribution of the data.
- To know about standardization
- To know about the central limit theorem

## TASKS

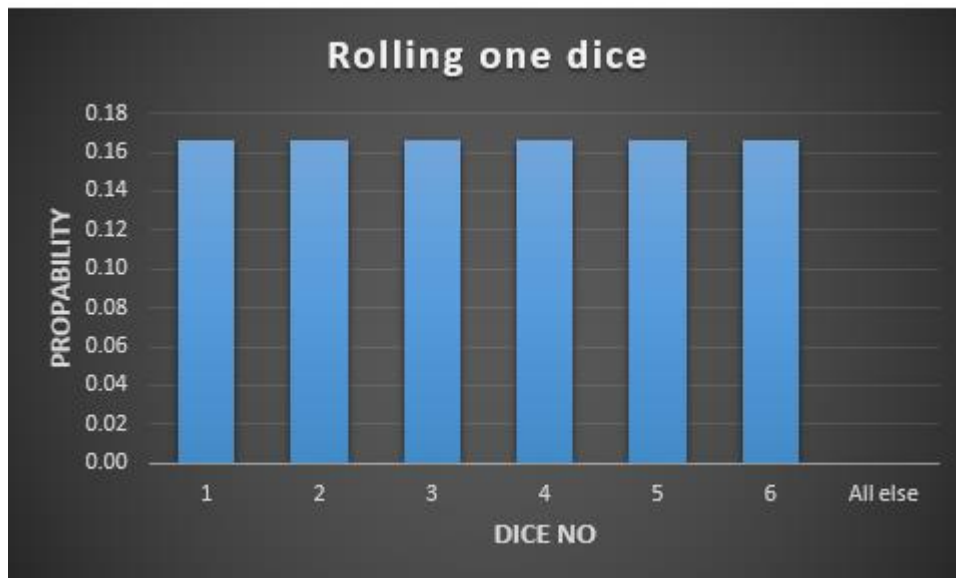
### Task 1:

#### Distribution:

In probability and statistics, a distribution is a mathematical function that gives the probabilities of occurrence of different outcomes for an experiment.

Outcome occurrence	
1	0.17
2	0.17
3	0.17
4	0.17
5	0.17
6	0.17
All else	0.00
1.00	

We roll a dice each no from 0 to 6 has equal probability of occurrence that is 0.17 and all other no has probability of zero.



Sum occurrence	
2	0.03
3	0.06
4	0.08
5	0.11
6	0.14
7	0.17
8	0.14
9	0.11
10	0.08
11	0.06
12	0.03
All else	0.00
1.00	



Rolling two dice. There is no chance of 1 to come because here we are using two dice and each dice has a minimum value of 1.

### Task 2:

Calculate the mean and standard deviation of the dataset. Standardize the dataset.

Mean		STD		
743.03		73.95306055		

Original dataset

567.45  
572.45  
572.45  
589.12  
613.87  
615.78  
628.45  
644.87  
650.45  
652.20  
656.87  
661.45  
666.45  
667.70  
668.95  
675.28  
675.78  
685.53  
694.28  
697.62  
705.78  
705.87  
708.12  
711.03

Standardization

-2.37417  
-2.30656  
-2.30656  
-2.08119  
-1.74652  
-1.7206  
-1.54932  
-1.32733  
-1.25184  
-1.22817  
-1.16507  
-1.10309  
-1.03548  
-1.01858  
-1.00168  
-0.91604  
-0.90928  
-0.77744  
-0.65912  
-0.61404  
-0.50361  
-0.50249  
-0.47206  
-0.43262

**Task 3:**

Parameter Values	
mean	Rs 100,200
SD	11478.40613
n	30
sqr (n)	5.477225575
sd error	2095.660653
C val (Z)	1.65
lower	Rs 96,742.53
higher	Rs 103,658.21

**90% confidential interval [96742.53, 103658.21]**