**Name = Garima Mittal**

**Roll No. = 210380**

**Assignment – 1**

**(Q.1)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Panel number** | **Film thickness (µm)** | **Average roughness (µm)** | **Current (A)** |
| 1 | 2.7 | 0.26 | 1.13 |
| 2 | 4.2 | 0.31 | 1.51 |
| 3 | 5.1 | 0.29 | 1.67 |
| 4 | 3.1 | 0.35 | 1.23 |
| 5 | 2.5 | 0.21 | 1.38 |
| 6 | 2.9 | 0.33 | 1.13 |
| 7 | 3.7 | 0.39 | 1.68 |
| 8 | 4.1 | 0.29 | 1.57 |
| 9 | 2.7 | 0.18 | 1.44 |
| 10 | 3.6 | 0.21 | 1.58 |

**Q1 (a).**

**Mean:**

1. Film Thickness (X̅)= = 3.46

2. Average roughness (Y̅) = = 0.282

3.Current (Z̅) = = 1.432

**Median:**

First sort in ascending order (N=10, even)

Median = mean of two central numbers.

1.Film thickness (µm) = [2.7, 2.7, 2.5, 2.9, 3.1, 3.6, 3.7, 4.1, 4.2, 5.1]

Median = 3.35

2. Average roughness (µm) = [0.18, 0.21, 0.21, 0.26, 0.29, 0.29, 0.31, 0.33, 0.35, 0.39]

Median = = 0.29

3.Current (A) = [1.13, 1.13, 1.23, 1.38, 1.44, 1.51, 1.57, 1.58, 1.67, 1.68]

Median = = 1.475

**Variance:**

1.Film thickness = ∑ = 0.6365

2. Average roughness = ∑ = 0.0040

3.Current = ∑= 0.0391

**Standard Deviation (S) =**

1.Film thickness = 0.840

2. Average roughness = 0.067

3. Current = 0.208

**Q1(b)**

**Z-scores**

Average\_Roughness: [-1.598, -1.128, -1.128, -0.345, 0.125, 0.125, 0.439, 0.752, 1.065, 1.692]

Film\_Thickness: [-1.215, -0.962, -0.962, -0.709, -0.456, 0.177, 0.304, 0.81, 0.936, 2.075]

Current: [-1.527, 0.394, 1.203, -1.021, -0.263, -1.527, 1.254, 0.698, 0.04, 0.748]

**Q1(C)**

**Kurtosis**

Average\_Roughness: -0.06 <3

Film\_Thickness: 0.88 <3

Current: -0.60<3

➢𝜅 = 3: Mesokurtic. Normal distribution

➢𝜅 < 3: Platykurtic

➢ 𝜅 > 3: Leptokurtic

Covariance between Average\_Roughness and Film\_Thickness: 0.019

Covariance between Average\_Roughness and Current: 0.0

Covariance between Film\_Thickness and Current: 0.126

Correlation between Average\_Roughness and Film\_Thickness: 0.367

Correlation between Average\_Roughness and Current: 0.023

Correlation between Film\_Thickness and Current: 0.805

**Q1(d)**

**Film Thickness:** The mean and median are quite close, suggesting a fairly symmetric distribution. Either could be used, but the mean gives a slightly more precise central value.

**Average Roughness:** The median is a bit higher than the mean, indicating a possible slight skewness or outliers. Median might be a better choice here.

**Current:** The median is higher than the mean, suggesting a potential skew in the data. The median might be a better representation of the center in this case.

**Q1(e)**

**Confidence Interval (CI)**

CI = X̅ ± Z

1. Z = 1.96 for CI = 95%

CI (Film Thickness) = (2.939,3.981)

CI (Average Roughness) = (0.241,0.323)

CI (Current)= (1.303,1.561)

1. Z = 2.576 for CI = 99%

CI (Film Thickness) = (2.775,3.945)

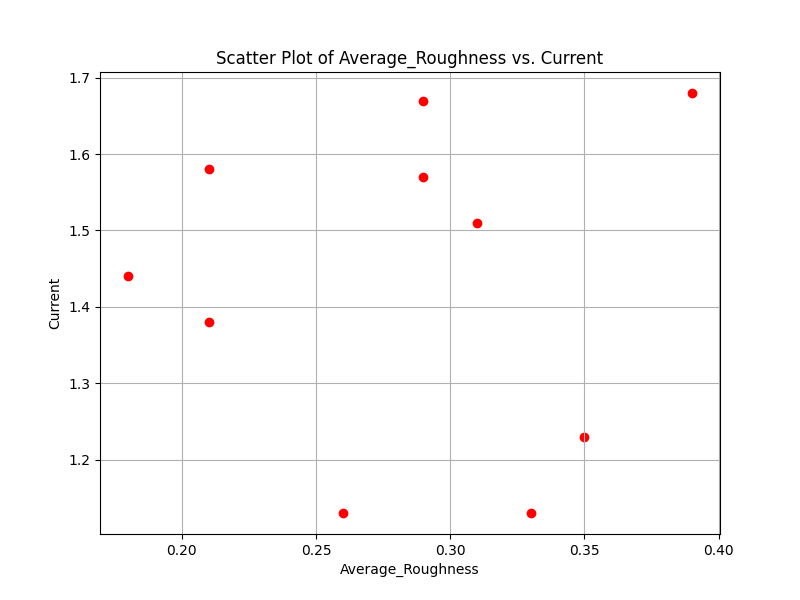
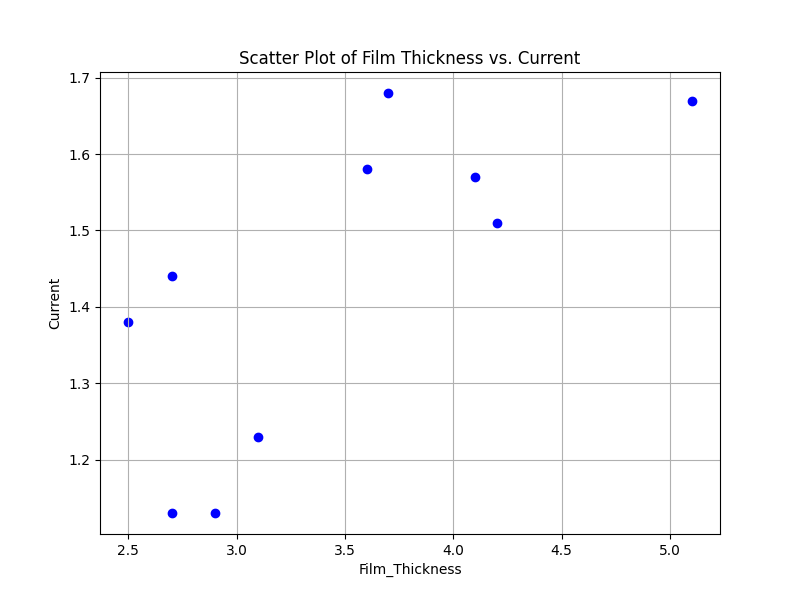
CI (Average Roughness) = (0.228,0.336)

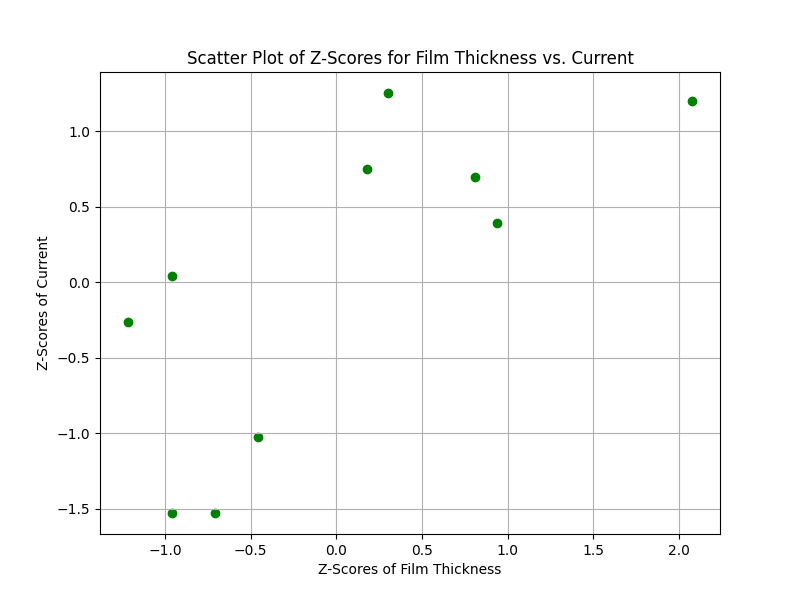
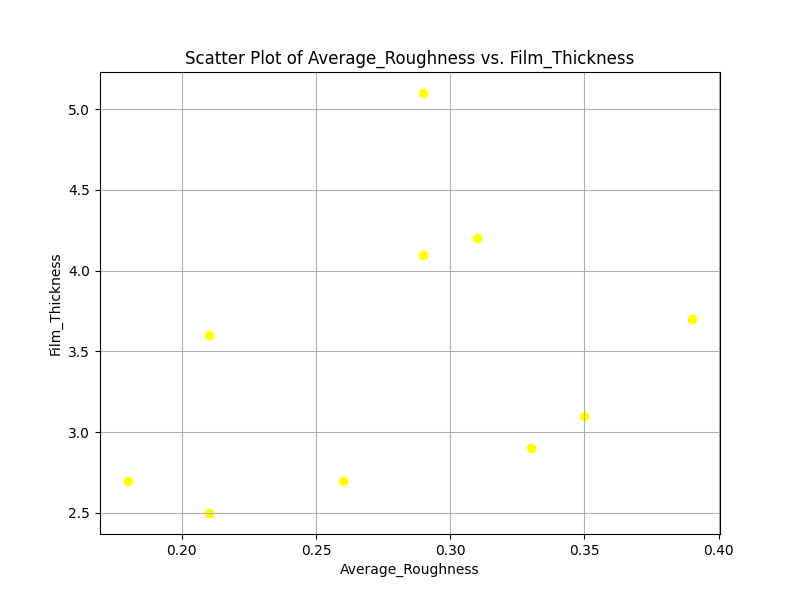
CI (Current)= (1.262,1.602**)**

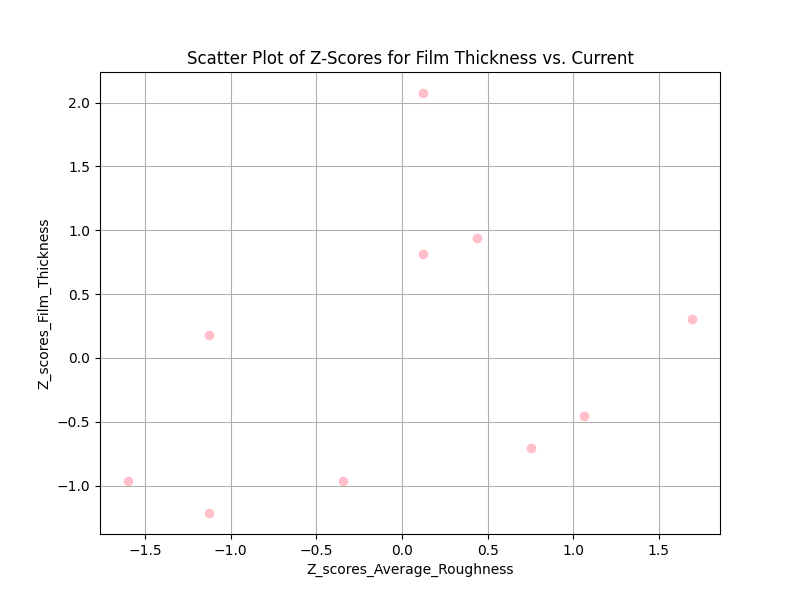
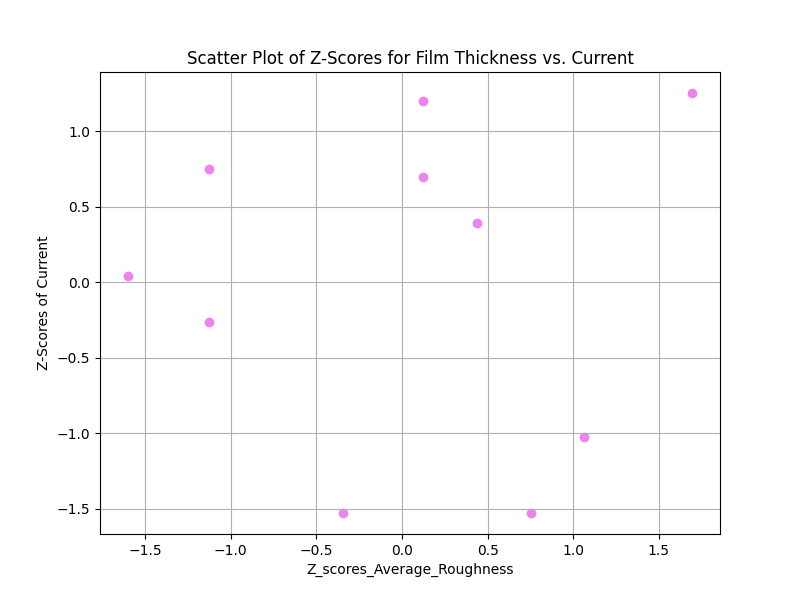
**Q1(f)**Generalization from the Batch

If this batch is representative of all batches, then:

* Descriptive Statistics**:** Mean, median, and other descriptive statistics might be generalizable to the entire production if the batch is representative.
* Confidence Intervals**:** The intervals calculated are based on this batch. If the batch is typical, these intervals would provide a good estimate of the parameters for the entire production. However, if this batch is not representative, these estimates might not accurately reflect the overall production.







**Q.2)**

|  |  |  |
| --- | --- | --- |
| Sample Number | Voltage(sample\_1) | Volatge(sample\_2) |
| 1 | 0.561 | 0.621 |
| 2 | 0.481 | 0.594 |
| 3 | 0.547 | 0.596 |
| 4 | 0.485 | 0.578 |
| 5 | 0.565 | 0.621 |
| 6 | 0.605 | 0.584 |
| 7 | 0.598 | 0.579 |
| 8 | 0.578 | 0.658 |
| 9 | 0.594 | 0.612 |
| 10 | 0.496 | 0.701 |

voltage is recorded at current of 2A/cm2

Mean:

Voltage(sample\_1) (X̅)= = 0.551

Voltage(sample\_2) (Y̅)= = 0.614

**Sample variance:**

= ∑ =

= 0.0022

= ∑ =

= 0.00153

**t-score(t) = = =** 74.3

**dof = =** = 46.73 = 47

**Q.3)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Country name** | **GDP Per capita** | **Social Support** | **Health Life Expectancy** |
| Afghanistan | 7.695 | 0.463 | 52.493 |
| Algeria | 9.342 | 0.802 | 66.005 |
| Belarus | 9.853 | 0.91 | 66.253 |
| Burundi | 6.635 | 0.49 | 53.4 |
| Chad | 7.364 | 0.619 | 48.478 |
| Colombia | 9.557 | 0.847 | 68.001 |
| El Salvador | 9.054 | 0.762 | 66.402 |
| Germany | 10.873 | 0.903 | 72.5 |
| India | 8.755 | 0.603 | 60.633 |
| Kuwait | 10.817 | 0.843 | 66.9 |
| Singapore | 11.488 | 0.915 | 76.953 |
| USA | 11.023 | 0.92 | 68.2 |
| Zambia | 8.145 | 0.708 | 55.809 |

1. **Mean GDP (X̅)= = 9.342**
2. **Mean\_ Social Support(Y̅)= =0.765**
3. **Mean\_ Health Life Expectancy (Z̅)= = 63.978**

**Standard Deviation (S)**

**1.Standard Deviation ()** = = 1.44

**2.Standard Deviation ()** = = 0.153

**3.Standard Deviation ()** = = 8.32

**Step 1.3: Standardize the Data**

The formula to standardize a data point is:

Standardized Value=

|  |  |  |  |
| --- | --- | --- | --- |
|  | GDP per capita | Social Support | Healthy Life Expectancy |
| 0 | -1.035 | -1.783 | -1.273 |
| 1 | 0.042 | 0.303 | 0.328 |
| 2 | 0.376 | 0.968 | 0.358 |
| 3 | -1.729 | -1.617 | -1.165 |
| 4 | -1.251 | -0.823 | -1.176 |
| 5 | 0.0183 | 0.580 | 0.565 |
| 6 | -0.145 | 0.057 | 0.375 |
| 7 | 1.044 | 0.925 | 1.098 |
| 8 | -0.341 | -0.921 | -0.308 |
| 9 | 1.007 | 0.556 | 0.434 |
| 10 | 1.447 | 0.999 | 1.626 |
| 11 | 1.142 | 1.032 | 0.588 |
| 12 | -0.740 | -0.275 | -0.880 |

Cov(X,Y) =

1

= 0.901

= 0.924

=1

= 0.867

=1

Covariance Matrix(**C**) =

[C-λ I] e = 0

[C-λ I] = 0

Det [C-λ I] = - = 0

On solving

**Eigenvalues:**

= 2.796

= 0.134

= 0.069

Eigenvectors:

**Eigenvectors:**

[[ 0.583745 0.16217166 0.7955766 ]

[ 0.57162828 -0.77794572 -0.26084779]

[ 0.57661329 0.60704267 -0.54682383]]

**Standardized Data:**

[[-0.77110291 -0.28645259 -6.79895332]]

**Projections on PCA1**:

[-1.77948434]

**Projections on PCA2:**

[4.92405004]

**Projections on PCA3:**

[-4.41443037]