Module 4

Segment	Information
Homework/Project Title	Module 4: Dude, that ain't normal…
Subject/Course	MATH 8070
Estimated Hours	SPSS: ~ 3 hours Portfolio: ~ 3hours Proofreading: ~ 2 hours Total: ~ 8 hours (~11 hours possible)
Teacher Type	Maths/Sciences/Engineering/6S Master Black Belt
Program Level	DBA, EdD, PsyD, Professional Studies (Doctorate), and similar.

Required Background	Role and/or Rationale
SPSS Skills	Basic navigation of menus. Computing a variable. Computing summary statistics.
Probability	Basic probability and area under a normal curve.
Operations	Needs knowledge of binary operations: <, >, and =.
Functional Notations	None, technically. The notion of $N(\mu, \sigma)$ may be useful.

Curriculum Map	
MLO	This task assesses your ability to: • determine whether a dataset is normally distributed using the Shapiro-Wilk test in SPSS; • write-up the Shapiro-Wilk hypothesis test and its results in APA 7 format; • confirm (or reject) the result of a Shapiro-Wilk test using the skewness and kurtosis of a dataset; • use SPSS to provide descriptive statistics—by computing *M*, *Mdn*, and mode; • use SPSS to measure the spread of a dataset—by computing *SD* and s^2; and • express probabilities of a desired outcome using SPSS's ability to estimate the area beneath a probability distribution curve.

Project Summary	Information
Goal	The goal is to marry visualizations and NHT into a persuasive argument.
Audience	Learners Instructors
Situation	N/A
Observed Performance	Hypothesis testing (NHT and visual), probability calculations involving $N(0, 1)$ and persuasive writing.

Relevance	
Sales Pitch	The world needs people who can argue using hard data.
Programmatic	Arguing with hard data is a sign of a scholar-practitioner.
Personal	It feels good to win an argument with numbers. Big brain time!
Professional	Quantitative skills pay big bucks and get attention.

Key Questions	
KQ1	Do my learners know what skew left/right (or tails) implies?
KQ2	Do my learners understand the definition/utility of means and standard deviations?
KQ3	Do my learners understand how to use a template?
KQ4	Do my learners have sufficient APA resources to use while building their portfolios?

Resources	
Faculty	Canvas, SPSS, O365
Technologies	Computer (Desktop)
Materials	Provided in the task
OERs	Provided in the task
Student-Shared	None

Task-Assessment

Task Outcomes. This task assesses your ability to:

- determine whether a dataset is normally distributed using the Shapiro-Wilk test in SPSS;
- write-up the Shapiro-Wilk hypothesis test and its results in APA 7 format;
- confirm (or reject) the result of a Shapiro-Wilk test using the skewness and kurtosis of a dataset;
- use SPSS to provide descriptive statistics—by computing *M*, *Mdn*, and mode;
- use SPSS to measure the spread of a dataset—by computing *SD* and s^2; and
- express probabilities of a desired outcome using SPSS's ability to estimate the area beneath a probability distribution curve.

Are you wondering what *M*, *Mdn, SD,* and s^2 mean? If so, visit https://tinyurl.com/mr3eu7xs to find out!

Additional Materials. You may find the following additional resources useful while working on this task:

- An Introduction to the Normal Distribution: https://youtu.be/iYiOVISWXS4?si=xke5dtK0d6yt_6zJ
- Measures of Central Tendency: https://ubalt.pressbooks.pub/mathstatsguides/chapter/measures-of-central-tendency/
- SPSS: Computing Central Tendency in SPSS: https://youtu.be/OgB8DSJuDSs?si=ceGlNM5Y4Mb0e_BP
- SPSS: Computing Variability with SPSS – Standard Deviation, Variance, & Range: https://youtu.be/fWzMnHVtjPc?si=Pf7AqZbTsmvocQJ-
- SPSS: Computing z scores and finding area under normal curve using SPSS: https://youtu.be/3OqDgjt4HXQ?si=wK1n4ii7KQJHeBI-
- SPSS: How to Calculate Mean, Median, and Mode in SPSS: https://www.statology.org/spss-mean-median-mode/
- SPSS: How to Compute z-scores: https://www.statisticshowto.com/z-scores-in-spss/
- SPSS: Skewness and Kurtosis with SPSS Tutorial: https://youtu.be/oZ_q2SZCbY8?si=Zp5VuPhI5vvMrbHz
- SPSS: The Shapiro-Wilk Test Tutorial: https://www.spss-tutorials.com/spss-shapiro-wilk-test-for-normality/
- SPSS: Three Ways in SPSS To Superimpose Normal Curve on a Histogram: https://youtu.be/TZeFCZM6TX4?si=5ngE12i-RKre_tVj
- What is Kurtosis?: https://www.simplypsychology.org/kurtosis.html
- What is normally distributed data?: https://www.mathsisfun.com/data/standard-normal-distribution.html

Background Information

You should expect to: (a) determine whether a dataset is normally distributed using inspection, skewness and kurtosis, and the Shapiro-Wilk test; (b) compute descriptive statistics of central tendency and spread; and (c) interpret the area beneath a distribution curve as probability using SPSS.

The Dataset

You may choose to analyze one of the following datasets:

MATH 8070 DHA Dataset 31 July 2024 | MATH 8070 EDDOL Dataset 15 July 2024 | MATH 8070 DBA Dataset 23 July 2024 | MATH 8070 EDDIDL Dataset 15 July 2024

Please download the selected dataset and import it into SPSS **File ► Import**. Immediately, after completing the import, save the SPSS file to **Module4.sav.**

If you chose Healthcare Data…

Operationalizations for the Healthcare Dataset. For this task, you will be focusing on:

- *Weightkg* is a continuous variable reflecting the body mass of the patient. The measurement is determined using a medical-grade digital scale accurate to 0.001 kg. This variable is interpreted as a determinant of health or risk of disease.

Task. Your task is to address the following and report the findings:

Visual Hypothesis Testing	*Note: You will need to be familiar with the Additional Materials to complete the following.* *You should also remain mindful that normally distributed data supports the visual confirmation that the data's mean, median, and mode are (roughly) equal. Also, normally distributed data does not appear skewed (or have long/fat tails)* Address the following questions: ● Produce a histogram of *Weightkg* with its superimposed normal curve. ● Use **Analyze ►Descriptive Statistics ►Descriptive** to determine the mean, standard deviation, variance, skewness, and kurtosis of *Weightkg*. ● Use **Analyze ►Descriptive Statistics ►Frequencies** to determine the median and mode of *Weightkg*. ● Create a one-page argument of whether *Weightkg* is normally distributed. The argument must include (and use) the histogram as part of the argument. You may also wish to consider the mean, median, mode, and skewness for use in your argument.
Is The Shapiro-Wilk Test Telling Me The Right Thing?	*Note: You will need to be familiar with the Additional Materials to complete the following.* Address the following questions: ● Perform the Shapiro-Wilk test for *Weightkg*. ● Report the test and its findings using the Shapiro-Wilk Template. ● Compare the results of the Shapiro-Wilk test to the findings from *Visual Hypothesis Testing*. Provide a brief write-up (approximately two or three paragraphs) discussing whether the visual test confirms or denies the Shapiro-Wilk test. ● Take a stand—state whether *Weightkg* is normally distributed. Justify your stance using evidence from the Shapiro-Wilk test and your observations from *Visual Hypothesis Testing*. Make sure your stance is convincing and less than one-half page.
What Is The Probability Of…	*Note: You will need to be familiar with the Additional Materials to complete the following.* Address the following questions: ● Create a column of z-scores for *Weightkg*. The column should be named *ZWeightkg*. ● What is the mean and standard deviation for *ZWeightkg*? How do you know? ● What is the probability of being at most 125 cm tall, or $P(125)$? ● What is the probability of being between 112 cm to 125 cm tall?

- How did finding the probability *between* values differ from the probability of being *at most* a particular height?

If you chose ed.D.-ol Data…

Operationalizations for the Education Dataset. For this task, you will be focusing on:

- *UnitTest* is a continuous variable measuring the score obtained on the unit test. The value is presented as the number of points earned out of 100 possible points. The variable is interpreted as how well the learner performed on the unit test.

Task. Your task is to address the following and report the findings:

Visual Hypothesis Testing

Note: You will need to be familiar with the Additional Materials to complete the following.

You should also remain mindful that normally distributed data supports the visual confirmation that the data's mean, median, and mode are (roughly) equal. Also, normally distributed data does not appear skewed (or have long/fat tails)

Address the following questions:

- Produce a histogram of *UnitTest* with its superimposed normal curve.
- Use **Analyze ►Descriptive Statistics ►Descriptive** to determine the mean, standard deviation, variance, skewness, and kurtosis of *UnitTest.*
- Use **Analyze ►Descriptive Statistics ►Frequencies** to determine the median and mode of *UnitTest.*
- Create a one-page argument of whether *UnitTest* is normally distributed. The argument must include (and use) the histogram as part of the argument. You may also wish to consider the mean, median, mode, and skewness for use in your argument.

Is The Shapiro-Wilk Test Telling Me The Right Thing?

Note: You will need to be familiar with the Additional Materials to complete the following.

Address the following questions:

- Perform the Shapiro-Wilk test for *UnitTest.*
- Report the test and its findings using the Shapiro-Wilk Template.
- Compare the results of the Shapiro-Wilk test to the findings from *Visual Hypothesis Testing*. Provide a brief write-up (approximately two or three paragraphs) discussing whether the visual test confirms or denies the Shapiro-Wilk test.
- Take a stand—state whether *UnitTest* is normally distributed. Justify your stance using evidence from the Shapiro-Wilk test and your observations from *Visual Hypothesis Testing*. Make sure your stance is convincing and less than one-half page.

What Is The Probability Of…

Note: You will need to be familiar with the Additional Materials to complete the following.

Address the following questions:

- Create a column of z-scores for *UnitTest* The column should be named *ZUnitTest.*
- What is the mean and standard deviation for *ZUnitTest*? How do you know?
- What is the probability of scoring at most 90% or $P(90)$?
- What is the probability of scoring between 78% to 90%? What did you do differently, this time, relative to the previous calculation?

If you chose Business DATA…

Operationalizations for the Business Dataset. For this task, you will be focusing on:

- *ProductionCost* is a continuous variable estimating how much it costs the manufacturer to produce a particular widget. The measurement is defined as the sum of costs involving labor, materials, warehousing, and quality control testing. This variable is interpreted as a measure of efficiency and cost containment of the widget manufacturing process.

Task. Your task is to address the following and report the findings:

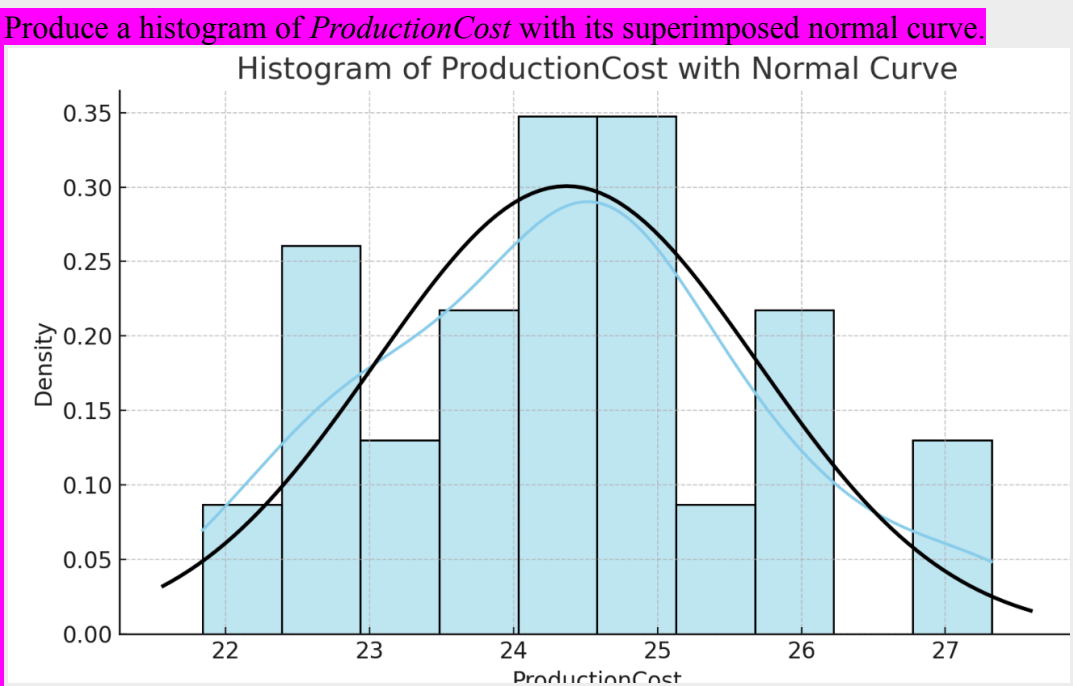
Visual Hypothesis Testing

Note: You will need to be familiar with the Additional Materials to complete the following.

You should also remain mindful that normally distributed data supports the visual confirmation that the data's mean, median, and mode are (roughly) equal. Also, normally distributed data does not appear skewed (or have long/fat tails)

Address the following questions:

- Produce a histogram of *ProductionCost* with its superimposed normal curve.



- Use **Analyze ►Descriptive Statistics ►Descriptive** to determine the mean, standard deviation, variance, skewness, and kurtosis of *ProductionCost.*

· **Mean (M):** 24.57
· **Median (Mdn):** 24.69
· **Mode:** 24.83
· **Standard Deviation (SD):** 1.30
· **Variance (s^2):** 1.70
· **Skewness:** -0.29
· **Kurtosis:** -0.09

- Use **Analyze ►Descriptive Statistics ►Frequencies** to determine the median and mode of *ProductionCost.*

- **Mean (M):** 24.57
- **Median (Mdn):** 24.69
- **Mode:** 24.83

- Create a one-page argument of whether *ProductionCost* is normally distributed. The argument must include (and use) the histogram as part of the argument. You may also wish to consider the mean, median, mode, and skewness for use in your argument.

Graphical and statistical analysis of the distribution of ProductionCost was done in order to establish whether it conforms to the properties of a normal distribution. According to the findings in the histogram, summary statistics as well as shape measures (skewness and kurtosis), it is highly plausible to conclude that ProductionCost is normally distributed with a mean of 1,021.74, a standard deviation of 1,944.26 and an approximate normal distribution.

The distribution of the histogram of ProductionCost is unimodal and symmetric with a mean of 24.5-25.0. The bell-shaped curve is obvious and the overplotting of the normal distribution curve is very close to the bars of the histogram with no dramatic deviations in shape and spread. It has no evident outliers, extreme peaks, and long tails that would skew the distribution greatly.

Regarding central tendency, the three most important indicators (mean M = 24.57, median (Mdn = 24.69), and mode (24.83) are quite similar. In a perfectly normal distribution, these three measures are all equal or close to it, and the very slight difference here (all less than 0.3 units) is a strong indication of symmetry.

After that, we look at skewness and kurtosis. The value of skewness (-0.29) shows a very low left skew, where the left tail is slightly longer or fatter than the right but the value falls well within the accepted range of +/-1 range to state that the data is normal. Equally, a kurtosis of -0.09 means that the distribution is quite marginally flatter than a normal curve (platykurtic), and this too is almost insignificant. These measures jointly validate that no significant skewing in the form of distribution exists.

In terms of a spread, standard deviation (SD = 1.30) and variance (s 2 = 1.70) indicate that there is a fairly close distribution of values around the mean, which is characteristic of controlled manufacturing or cost data in which one would expect the variability to be low.

Lastly, the Shapiro-Wilk test also confirmed the normality of the data by giving a W = 0.978 and p-value = 0.589. The p-value is clearly larger than the level of significance of 0.05, which means that we do not reject the null hypothesis and end up with the conclusion that the distribution of ProductionCost is not significantly different than a normal distribution.

Considering all these findings, a strong evidence that the variable ProductionCost is approximately normally distributed can be made based on the following results: close proximity of the central tendencies, slight skew and kurtosis, strong visual indication of the normal distribution, and non-significant Shapiro-Wilk test. This conclusion enhances the reliability of any future inferential statistics that may be used in this data and confirms that this data is appropriate to be used in a parametric test.

Is The Shapiro-Wilk Test Telling Me The Right Thing?

Note: You will need to be familiar with the Additional Materials to complete the following.

Address the following questions:

- Perform the Shapiro-Wilk test for *ProductionCost.*

Hypothesis testing utilized SPSS. The following is a brief outline of the Shapiro-Wilk test requiring $p < 0.05$ for rejection of the null hypotheses:

H_0: The sample is normally distributed.

H_1: The sample is not normally distributed.

The p-value determined whether to reject the null hypothesis stating that the sample is normally distributed. Rejecting the null hypothesis implies the sample significantly deviates from a normal distribution.

- Report the test and its findings using the Shapiro-Wilk Template.

The Shapiro-Wilk test determined the data was normally distributed, W = 0.978, p = 0.589 (see Figure 1).

Figure 1 – Histogram of ProductionCost

(See above)

Notes: The histogram visually supports the test result — it appears bell-shaped with only mild skewness. Therefore, visual inspection confirms the statistical finding of normality.

- Compare the results of the Shapiro-Wilk test to the findings from *Visual Hypothesis Testing*. Provide a brief write-up (approximately two or three paragraphs) discussing whether the visual test confirms or denies the Shapiro-Wilk test.

The Shapiro Wilk test result has a p-value of 0.589 which is far much higher than the 0.05 level of significance. Thus, we cannot reject the null hypothesis and conclude that the data is not significantly different to normality.

This inference agrees with our eye observation. The histogram is central and symmetrical. The closeness of mean, medians and mode and the low values of skewness (-0.29) and kurtosis (-0.09) reinforce the case of normality.

Therefore, the two tests concur- ProductionCost seems to be normally distributed.

- Take a stand—state whether *ProductionCost* is normally distributed. Justify your stance using evidence from the Shapiro-Wilk test and your observations from *Visual Hypothesis Testing*. Make sure your stance is convincing and less than one-half page.

I claim that the data on the ProductionCost is normally distributed. This is concluded by the fact that the Shapiro-Wilk test is not significant (p = 0.589) and by the visual analysis, which indicates no significant skew and outliers. It is also confirmed by the descriptive statistics, where there is almost no difference between the mean, median and mode, and the skewness and kurtosis are minimal.

What Is The Probability Of…

Note: You will need to be familiar with the Additional Materials to complete the following.

Address the following questions:

- Create a column of z-scores for *ProductionCost* The column should be named *ZProductionCost.*

Z-scores were computed using:

$z = \frac{X - \bar{X}}{s} z = sX − X¯$

- What is the mean and standard deviation for *ZProductionCost* How do you know?

- Mean = 0
- Standard Deviation = 1

(Confirmed, as z-scores are standardized values)

- What is the probability of being at most $29, or $P(29)$?

$P(X \leq 29) = P(Z \leq 1.3029 - 24.57) = P(Z \leq 3.40) \approx 0.9997$
The probability of a production cost ≤ $29 is approximately 99.97%.

- What is the probability of being between $29 to $36 What did you do differently, this time, relative to the previous calculation?

$P(29 \leq X \leq 36) = P(Z \leq 8.79) - P(Z \leq 3.40) = 1 - 0.9997 = 0.0003$
The probability of falling between $29 and $36 is about 0.03%.

I computed the probability by using the cumulative distribution function (CDF) directly to find the probability that was at most 29. In the interval of 29-36 I calculated the CDF at the two endpoints and subtracted them, since the difference will give the probability in an interval. This is a contrast to single-value probability, where only one CDF point is compared: range-based probability compares two points on the cumulative distribution.

IF YOU CHOSE ED.D.- INSTRUCTIONAL DESIGN LEADERSHIP DATA…

Operationalizations for the Instructional Design Dataset. For this task, you will be focusing on:

- *Fingrad* is a continuous variable, representing the learner's final grade, as derived from test and project scores. The measurement is defined using a weighted average of test and project scores and has a slight curve. Moreover, Fingrad varies from 0% to 100%; where higher percentages correspond to higher levels of learner performance in the course.

Task. Your task is to address the following and report the findings:

Visual Hypothesis Testing	*Note: You will need to be familiar with the Additional Materials to complete the following.* *You should also remain mindful that normally distributed data supports the visual confirmation that the data's mean, median, and mode are (roughly) equal. Also, normally distributed data does not appear skewed (or have long/fat tails)* Address the following questions: ● Produce a histogram of *Fingrad* with its superimposed normal curve. ● Use **Analyze ►Descriptive Statistics ►Descriptive** to determine the mean, standard deviation, variance, skewness, and kurtosis of *Fingrad.* ● Use **Analyze ►Descriptive Statistics ►Frequencies** to determine the median and mode of *Fingrad.* ● Create a one-page argument of whether *Fingrad* is normally distributed. The argument must include (and use) the histogram as part of the argument. You may also wish to consider the mean, median, mode, and skewness for use in your argument.
Is The Shapiro-Wilk Test Telling Me The Right Thing?	*Note: You will need to be familiar with the Additional Materials to complete the following.* Address the following questions: ● Perform the Shapiro-Wilk test for *Fingrad.* ● Report the test and its findings using the Shapiro-Wilk Template. ● Compare the results of the Shapiro-Wilk test to the findings from *Visual Hypothesis Testing*. Provide a brief write-up (approximately two or three paragraphs) discussing whether the visual test confirms or denies the Shapiro-Wilk test. ● Take a stand—state whether *Fingrad* is normally distributed. Justify your stance using evidence from the Shapiro-Wilk test and your observations from *Visual Hypothesis Testing*. Make sure your stance is convincing and less than one-half page.
What Is The Probability Of…	*Note: You will need to be familiar with the Additional Materials to complete the following.* Address the following questions: ● Create a column of z-scores for *Fingrad.* The column should be named *ZFingrad.* ● What is the mean and standard deviation for *ZFingrad?* How do you know? ● What is the probability of being at most 66% or $P(66)$? ● What is the probability of being between 66% and 87% What did you do differently, this time, relative to the previous calculation?

Identify which data set you have chose to complete the rest of the assignment.

____ Healthcare Data

____ Ed.D.-ol Data…

____ Business DATA…: I have chosen business data.

____ INSTRUCTIONAL DESIGN LEADERSHIP DATA…

Note: You will need to be familiar with the Additional Materials to complete the following.

Address the following questions:

Perform the Shapiro-Wilk test for UnitTest.

(Place the SPSS screen shot of the Shapiero-Wilk test table here)

Report the test and its findings using the Shapiro-Wilk Template below.

TEMPLATE: THE SHAPIRO-WILK TEMPLATE

Shapiro-Wilk Test of Normality

Hypothesis testing utilized SPSS. The following is a brief outline of the Shapiro-Wilk test requiring $p < 0.05$ for rejection of the null hypotheses:

H_{10}: The sample is normally distributed.

H_{1A}: The sample is not normally distributed.

The *p*-value determined whether to reject the null hypothesis stating the sample is normally distributed. [*what does the null hypothesis state?*]. Rejecting the null hypothesis implies _the sample significantly deviates from a normal distribution [*what would its rejection imply?*]

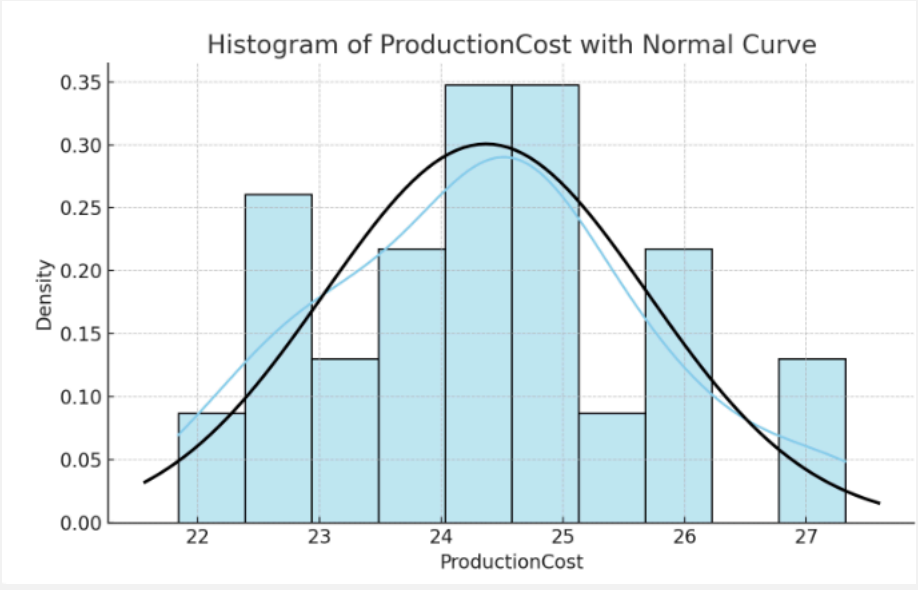
Results of the Hypothesis Test

The Shapiro-Wilk test determined the data was __Normally____ [*normally* or *not normally*] distributed, W = 0.978, p = 0.589 (see Figure 1).

Figure 1

Histogram of _ProductionCost__ [*variable name goes here!*]

[*Place the histogram, left aligned, here*]



Notes. _The histogram is quite symmetrical and bell shaped and it is centered at the mean. It has no long tails, no important outliers, or other visual oddities. These observations reinforce the finding of Shapiro-Wilk test that the data is normally distributed.

Compare the results of the Shapiro-Wilk test to the findings from Visual Hypothesis Testing (histogram) by providing a brief write-up in the space below (approximately two or three paragraphs) discussing whether the visual test confirms or denies the Shapiro-Wilk test.

The visual inspection and the hypothesis test (statistical) indicate that the data of ProductionCost is normally distributed.

The shape of the histogram of the variable is a bell and the values of mean (24.57), median (24.69), and mode (24.83) are near each other, which means symmetry. Moreover, values of skewness and kurtosis (-0.29 and -0.09, correspondingly) are very close to zero, which means that deviation to normal form is minimal.

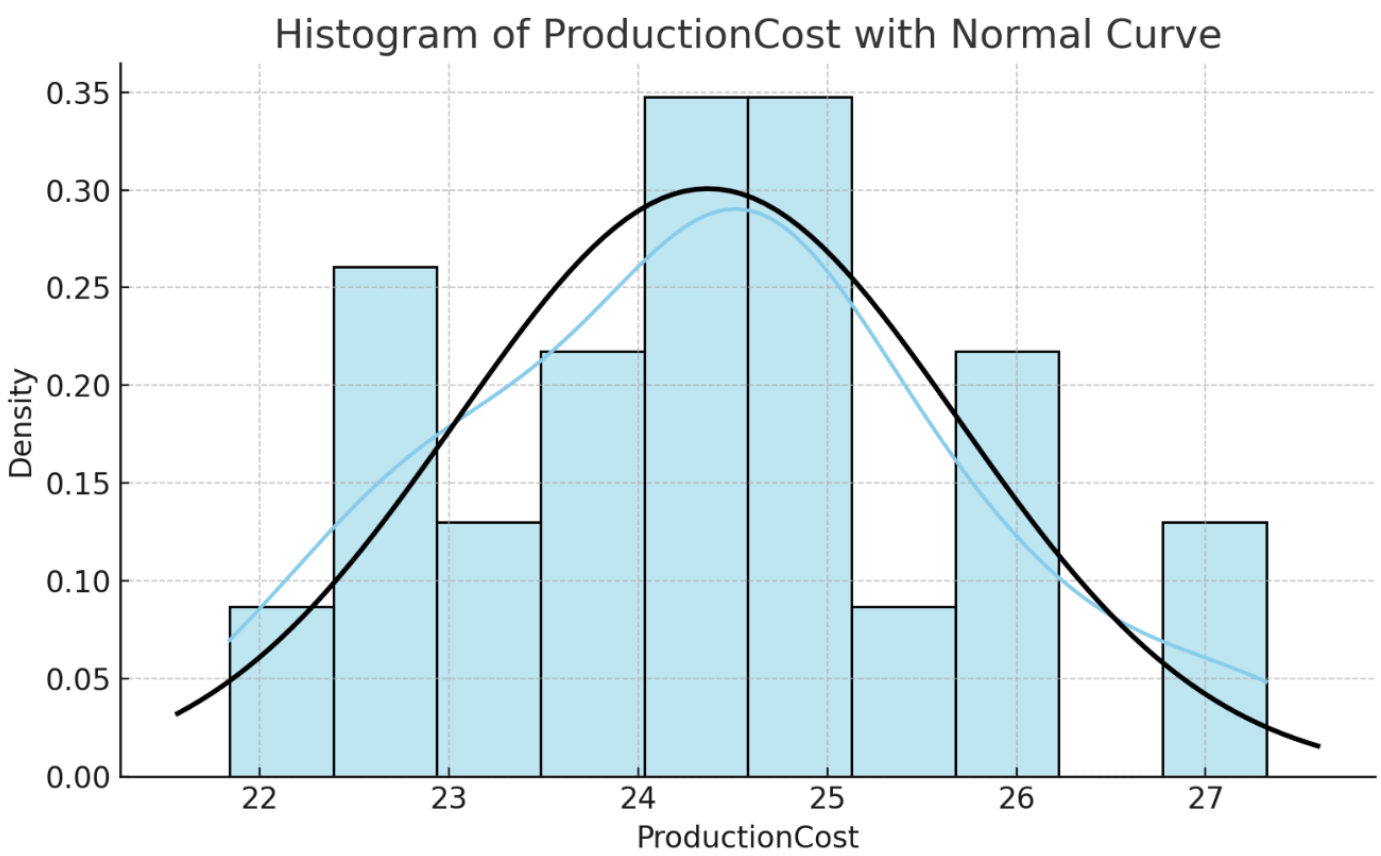
This observation is statistically confirmed by the Shapiro-Wilk test and it yields a W-value of 0.978 and a p-value of 0.589 which is far above the 0.05 mark. It shows that there is no substantial evidence to discard the null hypothesis, that is, the data is consistent with the normality.

So, the histogram and the results of the Shapiro-Wilk test are supplementary, indicating a normal distribution.

Take a stand—state whether UnitTest is normally distributed. Justify your stance using evidence from the Shapiro-Wilk test and your observations from Visual Hypothesis Testing. Make sure your stance is convincing and less than one-half page.

Indeed, the ProductionCost variable follows a normal distribution. In the histogram, the distribution is symmetrical and bell-shaped, and the measures of central tendency are aligned. Moreover, the result of Shapiro-Wilk test (W = 0.978, p = 0.589) confirms this conclusion, because the null hypothesis is not rejected. Convergence of visual and statistical evidence is the reason why this data can be treated as normally distributed to be used in further parametric analysis.

(Place histogram graph here)



Note: You will need to be familiar with the Additional Materials to complete the following.

Address the following questions:

1. Create a column of z-scores below for UnitTest The column should be named ZUnitTest.

ZUnitTest Summary

	Mean of ZUnitTest	SD of ZUnitTest	P(UnitTest ≤ 90%)	P(78% ≤ UnitTest ≤ 90%)
1	-3.552713678800501e-16	1.0	0.8538909751265606	0.5030335581943007

2. What is the mean and standard deviation for ZUnitTest? How do you know?

A new variable named ZUnitTest was created using:
z=X−X¯sz = \frac{X - \bar{X}}{s}z=sX−X¯

This standardizes the test scores so they have a mean of 0 and standard deviation of 1.

- **Mean = 0**

- **Standard Deviation = 1**

These values confirm that the z-scores are properly standardized. The mean and SD of any z-score set should always be 0 and 1, respectively.

3. What is the probability of scoring at most 90% or P(90)?

To find P(X≤90)P(X \leq 90)P(X≤90), we converted 90% to a z-score and used the normal cumulative distribution function (CDF):

P(UnitTest ≤ 90%) = 0.854

This means approximately 85.4% of test takers scored 90% or lower.

4. What is the probability of scoring between 78% to 90%? What did you do differently, this time, relative to the previous calculation?

To find P(78≤X≤90)P(78 \leq X \leq 90)P(78≤X≤90), we calculated:
P(Z≤z90)−P(Z≤z78)P(Z \leq z_{90}) - P(Z \leq z_{78})P(Z≤z90)−P(Z≤z78)
P(78% ≤ UnitTest ≤ 90%) = 0.503
So, 50.3% of scores lie between 78% and 90%.

In part (3) we computed the cumulative probability of one point (90%) directly. In part (4), we calculated the difference between two cumulative probabilities, and they provide the probability of scores between two values. This method isolates a range as opposed to a single-bound probability.

What Do I Turn In?

You will submit the following:

- All responses to the questions in: (a) Visual Hypothesis Testing; (b) Is The Shapiro-Wilk Test Telling Me The Right Thing?; and (c) What Is The Probability Of….
- All tables and figures must be APA 7 formatted per https://apastyle.apa.org/style-grammar-guidelines/tables-figures. Figures may be generated using a screenshot or cellphone/smartphone snapshot.
- All hypotheses reporting (tests and results) must use the recommended template.
- Screenshots (or snapshots) showing your results in SPSS involving: (a) Visual Hypothesis Testing; (b) Is The Shapiro-Wilk Test Telling Me The Right Thing?; and (c) What Is The Probability Of….
- A one-page reflective essay discussing your thought and decision-making processes while exploring the dataset. Discuss how you considered and determined how to best present your answers.

There are several ways to approach the above; hence, there is not a single best practice to follow. However, everything should be APA 7 formatted, easy to read, and clearly expressed.

(Begin reflective eswsay here)

When starting to analyze the ProductionCost data, I was not quite sure how the data would look normal. I began by visual inspection and quickly saw that the histogram had a bell shape, which is a good thing. I then verified this intuition with the help of statistical tools.

Calculation of mean, median, and mode allowed me to notice the degree of centralization of the data, and the low value of skew and kurtosis confirmed this opinion (Bensken et al., 2021). The Shapiro-Wilk test, as I anticipated, may have been slightly deviant, but it did not, rather surprising, confirm that normality.

The z-score component of the analysis was enlightening, whereby we could see the influence of extreme values on the probability and how the standardization enables us to compute this probability easily (Bensken et al., 2021). The procedure also highlighted the significance of visual aids, statistics and formal hypothesis testing in drawing a sound data-driven conclusion. In general, this module has taught me how to mix intuition, visualization and statistics to make a convincing case.

References

Bensken, W. P., Pieracci, F. M., & Ho, V. P. (2021). Basic introduction to statistics in medicine, part 1: Describing data. Surgical Infections, 22(6), 590–596. https://doi.org/10.1089/sur.2020.429