MODULE 5 UNIT 2

**ONGOING PROJECT**

**Learning outcomes:**

**LO3:** Use data visualisation to support analysis and interpretation of data.

**LO4:** Analyse data for pipelines and standard patterns in relation to broader industry contexts.

**LO5:** Compare data outputs to industry standards and benchmarks.

**LO6:** Formulate well-substantiated inferences based on statistical outcomes.

# Name:

## 1. Instructions and guidelines (Read carefully)

### Instructions

1. Insert your name and surname in the space provided above, as well as in the **file name.** Save the file as: **First name Surname Ongoing project (part 5)** – **e.g. Lilly Smith Ongoing project (part 5).** **(NB!** *Please ensure that you use the name that appears in your student profile on the Online Campus)*
2. Write all your answers in this document. There is an instruction that says, “Start writing here” under each question. Please type your answer there.
3. Submit your assignment in **Microsoft Word only**. No other file types will be accepted.
4. Do **not delete the plagiarism declaration** or the **assignment instructions and guidelines**. They must remain in your assignment when you submit. Do not remove or alter the template headers and footers.

**PLEASE NOTE:** **Plagiarism cases will be investigated in line with the Terms and Conditions for Students.**

**IMPORTANT NOTICE:** Please ensure that you have checked your course calendar for the due date for this assignment.

### Guidelines

1. There are 7 pages and 2 questions in this assignment.

2. Make sure that you have carefully read and fully understood the questions before answering them. Answer the questions fully but concisely and as directly as possible. Follow all specific instructions for individual questions (e. g. “list”, “in point form”).

3. You can access the raw data used in this project if you would like to run the analyses yourself, however this is not a prerequisite. All the information needed to answer the questions can be found in this document. To access the raw data, download the Excel workbook from the Ongoing project page on the Online Campus.

4. Answer all questions in your own words. Do not copy any text from the notes, readings or other sources. **The assignment must be your own work only.**

|  |
| --- |
| **Plagiarism declaration:** |
| **1. I know that plagiarism is wrong. Plagiarism is to use another’s work and pretend that it is one’s own.**  **2. This assignment is my own work.**  **3. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.**  **4. I acknowledge that copying someone else’s assignment (or part of it) is wrong, and declare that my assignments are my own work.** |

## 2. Mark allocation

Each question receives a mark allocation. However, you will only receive a final percentage mark and will not be given individual marks for each question. The mark allocation is there to show you the weighting and length of each question.

Question 1 9

Question 2 19

**TOTAL 28**

## 3. Questions

### Assignment instructions

#### Question 1

Sonic Selection would like to put firm performance metrics in place for its sales managers, based off customer conversion rates, as they would like to reward the sales managers based on the rate at which users that browse the site end up purchasing tracks. The need for changing the incentive scheme has been driven by information recently published by the eCommerce Association around industry conversion rates.

The eCommerce Association collects and analyses data from a broad range of companies involved with trading of goods or offering services over any electronic network, predominantly the internet. This incorporates all online stores, including digital music stores like Sonic Selection. The eCommerce Association has conducted a study on 312 e-commerce companies, which indicates that the expected conversion rate is 8%, with a standard deviation of 1.1%. This benchmark is far higher than Sonic Selection’s internal benchmark of a 5% conversion rate.

1.1 Given the recently published benchmarking information, Sonic Selection is well below the benchmark in terms of e-commerce conversion rates. What other factors, besides industry benchmarks, need to be considered when it comes to setting performance metrics? Name and briefly discuss at least 3 other factors. Your answer may not exceed the word limit.

(Max. 150 words)

Start writing here:

1.2 Based on the information provided, and your answer to question 1.1, suggest a suitable conversion rate for Sonic Selection and justify your choice. Should it be higher or lower than the industry average? Your answer may not exceed the word limit.

(Max. 100 words)

Start writing here:

#### Question 2

Sonic Selection would like to conduct A/B testing to improve the average revenue per site visitor. In order to reach their target, which is an average increase of $200.00 in revenue per site visitor, a new landing page is designed and then compared to the original website landing page. As the data analyst at Sonic Selection, you have been tasked with conducting an analysis as to whether the new website design achieves an increase in revenue per site visitor.

**A/B testing re-cap:**

If you would like to revise what A/B testing entails, follow this link to a brief video explaining [what A/B testing is](https://www.youtube.com/watch?v=9zAQw6UhQLs), or revisit the interactive video in Unit 1.

2.1 State your null and alternative hypotheses for the A/B testing research project, in the spaces provided.

* Null Hypothesis (H0)  
  Start writing here:
* Alternative hypothesis (HA)  
  Start writing here:

2.2 Figure 1 and Figure 2 display the distributions for the for the old website and the new website, respectively. The distribution of the data can be assessed by plotting the observations using histograms. The data is grouped into categories, or bins, to smooth the data and display the number of cases in each bin. When generating histograms in Excel, you can specify the bins. If the bins are not provided, Excel generates the bins automatically.

Referring to the histograms of the old and new website, comment on the distribution of the two samples (old website and new website). State whether the samples are equally distributed, or if the samples are positively or negatively skewed. Also include a brief explanation of how the graphs differ from one another, and what this may imply. Pay attention to the scales of the graphs. Your answer may not exceed the word limit.

(Max. 100 words)

**Figure 1:** Histogram of the distribution of the average revenue for the old website.

**Figure 2:** Histogram of the distribution of the average revenue for the new website.

Start writing here:

2.3 Using a 0.05 alpha level, an F-test has been conducted to check if the two samples have equal or unequal variances. Excel returns the following table:

|  |  |  |
| --- | --- | --- |
| **F-Test Two-Sample for Variances** | |  |
|  |  |  |
|  | *New Site Avg. Revenue Per*  *Visitor* | *Old Site Avg. Revenue Per Visitor* |
| Mean | $1 238,43 | $570,63 |
| Variance | $524743,33 | $207230,17 |
| Observations | 30 | 30 |
| df | 29 | 29 |
| F | 2.532176 |  |
| P(F<=f) one-tail | 0.0073850 |  |
| F Critical one-tail | 1.8608114 |  |

What conclusions can be drawn from the results of the F-test, and what are the implications of these results? Justify your answer by referring to the relevant values in the table above. Your answer may not exceed the word limit.

(Max. 100 words)

Start writing here:

2.4 A t-test has been conducted to determine whether the average revenue per site visitor has increased by the target of $200.00. Based on your answer in Question 2.3, do you think a two-sample t-test with equal variance or the two-sample t-test with unequal variance is appropriate? State your choice with your justification. Your answer may not exceed the word limit.

(Max. 50 words)

Start writing here:

2.5 After applying the t-tests in Excel, the following tables are produced:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **T-Test: Two-Sample Assuming Unequal Variances** | | | |  |
|  |  | | |  |
|  | *New Site Avg. Revenue Per Visitor* | | | *Old Site Avg. Revenue Per Visitor* |
| Mean | 1238,427015 | | | 570,634596 |
| Variance | 524743,3286 | | | 207230,1741 |
| Observations | 30 | | | 30 |
| Hypothesized Mean Difference | 200 | | |  |
| df | 49 | | |  |
| t Stat | 2.994788293 | | |  |
| P(T<=t) one-tail | 0.002148695 | | |  |
| t Critical one-tail | 1.676550893 | | |  |
| P(T<=t) two-tail | 0.004297391 | | |  |
| t Critical two-tail | 2.009575237 | | |  |
| **T-Test: Two-Sample Assuming Equal Variances** | | | | |
|  | |  |  | |
|  | | *New Site Avg. Revenue Per Visitor* | *Old Site Avg. Revenue Per Visitor* | |
| Mean | | 1238,427015 | 570,634596 | |
| Variance | | 524743,3286 | 207230,1741 | |
| Observations | | 30 | 30 | |
| Pooled Variance | | 365986,7514 |  | |
| Hypothesized Mean Difference | | 200 |  | |
| df | | 58 |  | |
| t Stat | | 2.994788293 |  | |
| P(T<=t) one-tail | | 0.002016711 |  | |
| t Critical one-tail | | 1.671552762 |  | |
| P(T<=t) two-tail | | 0.004033423 |  | |
| t Critical two-tail | | 2,001717484 |  | |

What can you conclude about the new version of the Sonic Selection website, and which version of the website would you suggest Sonic Selection use? Justify your answer by referring to the relevant values in the table above. Be sure to explicitly state how you reached your conclusion based on the interpretation of the relevant statistical values. In your concluding statement, refer back to your hypotheses, and state whether the null hypothesis can be accepted or rejected. Your answer may not exceed the word limit.

(Max. 200 words)

Start writing here: