**SCENARIO 1**

**EXPERIMENT DESIGN:**

**1.** **What are objects, subjects, treatments and factors in this experiment?**

**Answer:**

**Objects:** In the experiment, Program-1 and Program-2 are the objects.

**Subjects:** The 12 students were considered as subjects because, they were assigned to conduct the experiments on IDE’s, using 4 different designs.

**Treatment:** IDE-A or IDE-B are considered as treatments in the experiment.

**Factors:** The factor in the experiment is Efficiency, as it is measured to find which design is efficient.

**2. How would you describe Design-A and Design-D in terms of a standard design type e.g. one factor, two treatments?**

**Answer:** In Design-A, one program and two IDE’s (IDE-A and IDE-B) were taken. So, the one factor is a program and two treatments are testing the program using two IDE’s.

In Design-B, the two programs were tested using two IDE’s. So, in one case the program is considered as one factor and the IDE’s as two treatments whereas, in another case, IDE is considered as one factor and program as two treatments.

**3.** **What are the benefits and limitations of using Design-B instead of Design-A?**

**Answer:** The benefit of using Design-B is that, every person works on both IDE-A and IDE-B and so that every individual opinion can be considered while figuring out the efficient IDE.

Limitation: As every person performs on both the IDE’s, it is time to consume to complete the experiment.

**4.** **What problems/mistakes can you identify in Design-C?**

**Answer:** Every person is testing both the IDE’s with two different programs. By considering those values, one cannot conclude which IDE is efficient. To compare the efficiencies of both the IDE’s, they should be tested with any one program.

**5.** **Does Design-D solve the problems you have identified in Design-B and Design-C?**

**Answer:** Yes, the problem with Design-B is more consumption of time and that problem is solved in Design-D, by distributing the tasks. The problem in Design-C is using different programs for different IDE’s to find the efficiency and this problem is solved by Design-D by taking one program for both the IDE’s.

**6. What are the benefits and limitations of the designs Design-A and Design-D? E.g. one is easier to analyse.**

**Answer:** Benefits of Design-A is less time consuming and so the efficient IDE can be found quickly.

The result may not be accurate because each person tests only one IDE and cannot justify the efficient IDE without testing the other.

Benefits of Design-D are, both the IDE’s were tested with two different programs by assigning each probable case to 3 students. This gives the best result.

The problem with Design-D is the same person is not involved in doing the tests for two different IDE’s or same IDE with a different program.

**7.** **What variables must be controlled in Design-A to increase the validity of the experiment? E.g. previous experience/familiarity with subjects to IDE ’s?**

**Answer:** To increase the validity of the experiment the person who is familiar with any one of the IDE should be assigned to test the other IDE. By doing so, the results will be more accurate.

**EXPERIMENT ANALYSIS:**

**1.​State the null and alternative hypothesis for this investigation?**

**Answer:** ​Null hypothesis: H0: µIDE-A =µIDE-B

Alternative hypothesis: H1: µIDE-A ≠µIDE-B

**2.​ Use descriptive statistics and visualize the data in Table 7 uses e.g. box plot, histograms and scatter plot. Which visualization tool helped you develop some insights into the data? What were the insights e.g. any interesting patterns or trends in the data, a clear difference in efficiency between two IDEs, outliers?**

**Answer​:** With the help of Excel, we designed the scatter-plot with the values of IDE-A and IDE-B.

**Time consumed by IDE-A for program development**

**Time consumed by IDE-B for program development**

**3.​** **Choose and justify your choice of a parametric/non-parametric test for analysing the given data.**

**Answer:** When we calculate the Standard deviations of IDE-A and IDE-B. The results are as follows,

(i) SD of IDE-A: √∑(X-X1)2/N = 27.06

SD of IDE-B: √∑(X-X1)2/N = 21.24

The SD of IDE-A is greater than SD of IDE-B.

(ii) The confidence interval of IDE-A with 95% confidence is 19.76

The confidence interval of IDE-A with 99% confidence is 25.97

The confidence interval of IDE-B with 95% confidence is 15.52

The confidence interval of IDE-B with 99% confidence is 20.39

In both the cases, the confidence interval of IDE-A is greater than IDE-B.

The relation between confidence interval and confidence is “lesser the confidence interval, more the confidence on the values.”

As the deviation and confidence interval is more for IDE-A than IDE-B, IDE-B is said to be efficient.

**4.​ Run the statistical method and report if you can reject the null hypothesis? Please interpret your results, what does this imply for the objective of the study?**

**Answer: ​** According to the null hypothesis, mean values of two IDE’s should be related as µIDE-A =µIDE-B. As the mean values of both the IDE’s were not equal, Null hypothesis can be rejected.

**5.​ Based on the results would you be confident to recommend an IDE either IDE-A or IDE-B for use in your company. Why or why not?**

**Answer:** By considering the results of both the IDE’s, we recommend IDE-B rather than IDE-A as the deviation of IDE-B is less with the normal.

**SCENARIO - 2**

**a.) Describe the approach that you will follow to analyze the given data (i.e. the three**

**papers identified in Section 3.1.2). Please read Chapter 18 of C. Robson, K. McCartan,**

**Real world research: A resource for social scientists and practitioner-researchers.**

**Fourth Edition. Wiley, 2016, to make an informed decision about your approach and**

**the steps you take. For example, the analysis approach you will use (a. Quasistatistical**

**approach, b. thematic coding approach or c. grounded theory approach).**

**Also describe your mechanism for coding the data. Also explain why you chose the**

**approach over other alternatives.**

**Answer:** The approach is the Grounded Theory for data analysis. Grounded Theory

approach is preferred to the other approaches for its unique way of analysis through

theory type of data. This theory is taken from the references i.e., collection of data from

the sources and analyzed. The data is categorized and description of the core data is

identified by which the relationships are conceptualized. This mechanism is done in

three stages.

Finding the relative categories of the data.

Finding the relationships between the categories.

Conceptualize and derive the relationships through finding the core categories.

This approach is taken as a better approach than the other as all these approaches are

virtually indistinguishable i.e., not able to be identified different when compared.

**b.) Please describe the coding procedure that you followed. For each step, please provide**

**an example of how you coded the information in the papers (see Section 3.1.2 for the**

**list of papers).**

**Answer:** Grounded Theory approach has its own coding procedure, which is related as

its achievements. This coding procedure is taken into consideration. The procedure is

achieved in three kinds of coding. They are as follows

1. **OPEN CODING**: To find the categories.

**Eg**: The data presented from the resources are considered and the required part

of the data is taken and represented as category or categories.

2. **AXIAL CODING**: To interconnect categories through the relationship of data.

**Eg**: This helps the data to interpret, divided into parts by interconnecting the

categories i.e., combining the data and presenting it in a specific order.

3. **SELECTIVE CODING**: To establish the core category or categories.

**Eg**: A specific part of the data is taken as the only base or as the core for the

establishment of the code analysis.

**c.) Answer the following questions by citing examples from your analysis of the three**

**studies (provided in Section 3.1.2) :**

**1.) Which challenges or impediments for industry-academia collaborations have been**

**raised by the papers?**

**Ans: CHALLENGES:**

In industrial-academia, current environment based knowledge is taken into

consideration than the empirical based knowledge [2].

Searching for the perfect or consistent host for the research project to adjust

for different difficulties and time period [3] is involved in collaboration [4].

Collaboration aims and objectives i.e., [3] goals are to be identified.

Direct and effective interaction is necessary in the research with the other

partners or [3] members.

Collecting the data without any complete information [3] is quite difficult.

The developer’s availability [4] is limited when connecting for a project and

non-permanent developers have enough time to contribute for the research

is also challenging.

**2.) What patterns have been proposed for industry-academia collaborations?**

**Ans: PATTERNS:**

Continuous interactions [4] i.e., meetings among the representatives gives a

prospective explanation about the industry and also consistency on further

projects.

Representatives should get worried about the benefits of the industry more

than the research people [3].

Each and every side i.e., department of the industry should be inspected

consistently such as global advertising, marketing, sales [4] etc.

Consistency should be present between the industry and academia

collaboration [4].

Commitments, shares should be given a basic attention to maintain and

develop [4].

**3.) What should be avoided during industry-academia collaborations?**

**Ans: AVERTIONS:**

Re-organization leads to new establishments [2] of work processes and

developments [4].

Research projects which are not paid, funding may get pressured from the

industry to the academia [2].

**REFERNCES:**

[1] Chapter 8 and 10 of Wohlin et al., Experimentation in Software Engineering,

Springer-Verlag Berlin Heidelberg, 2012

[2] P. Runeson, “It takes two to tango - An experience report on industry - Academia

collaboration,” *Proc. - IEEE 5th Int. Conf. Softw. Testing, Verif. Validation, ICST 2012*, pp.

872–877, 2012.

[3] S. Martinez-Fernandez and H. M. Marques, “Practical experiences in designing and

conducting empirical studies in industry-academia collaboration,” *Proc. 2nd Int. Work.*

*Conduct. Empir. Stud. Ind. - CESI 2014*, pp. 15–20, 2014.

[4] T. Gorschek, C. Wohlin, P. Carre, and S. Larsson, “A Model for Technology Transfer in

Practice,” *Ieee Softw.*, vol. 23, no. 6, pp. 88–95, 2006.

[5] Chapter 18 of C. Robson, K. McCartan, Real world research: A resource for social

scientists and practitioner-researchers. Fourth Edition. Wiley, 2016.