NAME = ABDUL QADIR

ID NO = 14097

PAPER = DATA ANALYSIS USING SPSS

INSTRUCTOR = SIR RAZA AHMAD

Q NO. 1 (PART A)

Fundamental Steps For Research?

The following seven steps outline a simple and effective strategy for finding information for a research paper and documenting the sources you find. Depending on your topic and your familiarity with the library, you may need to rearrange or recycle these steps. Adapt this outline to your needs. We are ready to help you at every step in your research.

STEP 1: *IDENTIFY AND DEVELOP YOUR TOPIC*

SUMMARY: *State your topic as a question. For example, if you are interested in finding out about use of alcoholic beverages by college students, you might pose the question, "What effect does use of alcoholic beverages have on the health of college students?" Identify the main concepts or keywords in your questions.*

STEP 2: *FIND BACKGROUND INFORMATION*

SUMMARY: *Look up your keywords in the indexes to subject encyclopedias. Read articles in these encyclopedias to set the context for your research. Note any relevant items in the bibliographies at the end of the encyclopedia articles. Additional background information may be found in your lecture notes, textbooks, and more*

STEP 3: USE CATALOGS TO FIND BOOKS AND MEDIA

SUMMARY: Use guided keyword searching to find materials by topic or subject. Print or write down the citation (author, title,etc.) and the location information (call number and library). Note the circulation status. When you pull the book from the shelf, scan the bibliography for additional sources. Watch for book-length bibliographies and annual reviews on your subject; they list citations to hundreds of books and articles in one subject area.

STEP 4: USE INDEXES TO FIND PERIODICAL ARTICLES

SUMMARY: Use periodical indexes and abstracts to find citations to articles. The indexes and abstracts may be in print or computer-based formats or both. Choose the indexes and format best suited to your particular topic; ask at the reference desk if you need help figuring out which index and format will be best. You can find periodical articles by the article author, title, or keyword by using the periodical indexes in the Library home page. If the full text is not linked in the index you are using, write down the citation from the index and search for the title of the periodical in the Cornell Library Classic Catalog. The catalog lists the print, microform, and electronic versions of periodical indexes at Cornell.

STEP 5: FIND ADDITIONAL INTERNET RESOURCES

Nearly everyone is aware of and uses <u>Google</u> and its branches, <u>Google Scholar</u>, <u>Google Books</u>, <u>Google News</u>, <u>YouTube</u>, etc., to search and find information on the open Internet (as opposed to the subscription-only resources you will encounter in steps 2 through 4 above). Here are links to other <u>search engines</u>.

You can also check to see if there is a research guide (a <u>subject guide or a course guide</u>) created by librarians specifically for your topic or your class that links to recommended resources.

STEP 6: <u>EVALUATE WHAT YOU FIND</u>

SUMMARY: *See <u>How to Critically Analyze Information Sources</u> and <u>Distinguishing Scholarly from Non-Scholarly Periodicals: A Checklist of Criteria</u> for suggestions on evaluating the authority and quality of the books and articles you located.*
If you have found too many or too few sources, you may need to narrow or broaden your topic. Check with a reference librarian or your instructor.

When you're ready to write, here is <u>an annotated list of books</u> to help you organize, format, and write your paper.

STEP 7: CITE WHAT YOU FIND USING A STANDARD FORMAT

Give credit where credit is due; cite your sources.

Citing or documenting the sources used in your research serves two purposes, it gives proper credit to the authors of the materials used, and

it allows those who are reading your work to duplicate your research and locate the sources that you have listed as references.

Knowingly representing the work of others as your own is plagarism. (See Cornell's <u>Code of Academic Integrity</u>). Use one of the styles listed below or another style approved by your instructor. Handouts summarizing the APA and MLA styles are available at Uris and Olin Reference.

Available online*:*

RefWorks *is a web-based program that allows you to easily collect, manage, and organize bibliographic references by interfacing with databases. RefWorks also interfaces directly with Word, making it easy to import references and incorporate them into your writing, properly formatted according to the style of your choice*

<u>Q NO. 1 (PART B)</u>

<u>*Inductive vs. deductive reasoning.*</u>

*The main difference between inductive and deductive reasoning is that inductive reasoning aims at **developing a theory** while deductive reasoning aims at **testing an existing theory**.*

Inductive reasoning moves from specific observations to broad generalizations, and deductive reasoning the other way around.

Both approaches are used in various <u>types of research</u>, and it's not uncommon to combine them in one large study.

Inductive research approach

When there is little to no existing literature on a topic, it is common to perform inductive research because there is no theory to test. The inductive approach consists of three stages:

1. ### Observation
   - A low-cost airline flight is delayed
   - Dogs A and B have fleas
   - Elephants depend on water to exist
2. ### Observe a pattern
   - Another 20 flights from low-cost airlines are delayed
   - All observed dogs have fleas
   - All observed animals depend on water to exist
3. ### Develop a theory
   - Low cost airlines always have delays
   - All dogs have fleas
   - All biological life depends on water to exist

Limitations of an inductive approach

A conclusion drawn on the basis of an inductive method can never be proven, but it can be invalidated.

Example
You observe 1000 flights from low-cost airlines. All of them experience a delay, which is in line with your theory. However, you can never prove that flight 1001 will also be delayed. Still, the larger your dataset, the more reliable the conclusion.

Deductive research approach

When conducting deductive research, you always start with a theory (the result of inductive research). Reasoning deductively means testing

these theories. If there is no theory yet, you cannot conduct deductive research.

The deductive research approach consists of four stages:

1. ***Start with an existing theory***
   - *Low cost airlines always have delays*
   - *All dogs have fleas*
   - *All biological life depends on water to exist*
2. ***Formulate a hypothesis based on existing theory***
   - *If passengers fly with a low cost airline, then they will always experience delays*
   - *All pet dogs in my apartment building have fleas*
   - *All land mammals depend on water to exist*
3. ***Collect data to test the hypothesis***
   - *Collect flight data of low-cost airlines*
   - *Test all dogs in the building for fleas*
   - *Study all land mammal species to see if they depend on water*
4. ***Analyse the results: does the data reject or support the hypothesis?***
   - *5 out of 100 flights of low-cost airlines are not delayed = reject hypothesis*
   - *10 out of 20 dogs didn't have fleas = reject hypothesis*
   - *All land mammal species depend on water = support hypothesis*

Limitations of a deductive approach

The conclusions of deductive reasoning can only be true if all the premises set in the inductive study are true and the terms are clear.

Example

- *All dogs have fleas (premise)*

- *Benno is a dog (premise)*
- *Benno has fleas (conclusion)*

Based on the premises we have, the conclusion must be true. However, if the first premise turns out to be false, the conclusion that Benno has fleas cannot be relied upon

Q NO.2 (PART A):

VALIDITY OF DATA:

In general, VALIDITY is an indication of how sound your research is. More specifically, validity applies to both the design and the methods of your research. Validity in data collection means that your findings truly represent the phenomenon you are claiming to measure. Valid claims are solid claims. When moving and merging data it's important to make sure data from different sources and repositories will conform to business rules and not become corrupted due to inconsistencies in type or context. The goal is to create data that is consistent, accurate and complete so to prevent data loss and errors during a move. In data warehousing, data validation is often performed prior to the ETL (Extraction Translation Load) process. A data validation test is performed so that analyst can get insight into

the scope or nature of data conflicts. Data validation is a general term and can be performed on any type of data, however, including data within a single application (such as Microsoft Excel) or when merging simple data within a single data store.

DATA CHECKING:

After having entered the data, you should carefully check the data to ensure that they have been entered correctly.

Sometimes erroneous data input will become apparent when looking at the data range in the summary statistics report (e.g. maximum value of 78 for pH), or when plotting box-and-whisker plots, dot plots or scatter diagrams for the different variables. You should check clear outliers since they may indicate incorrect data entry, or they may result from a technical failure in measurement or from a study protocol violation. Only for such plausible reason you may exclude a value from further analysis, and not simply because a value is the smallest or largest. If there is no evidence of such a mistake then the value must remain unaltered.

You can locate any value in the spreadsheet using the Find procedure.

You can exclude outliers from further calculations by using the Exclude command.

Q NO. 2 (PART B

INTREPRETATION OF THE TABLE:

In this table the mean for the writing score was 52.7750 it defines that the data was collected for the writing score and it has been found that average of the writing score 52.7750. next in this table we have 95% interval and if we trimmed the data 5% then the 95% confidence interval mean we have upper and lower case and the writing score for 95% confidence interval mean lies between the mean of 51.4533 and 54.0967 and 95% doubt is being cleared and we will get the average between 51.4533 and 54.0967. then we have next 5% trimmed data in this descriptive data. In this case for writing score the data that is being trimmed 5% and the average we are getting is 53.1389. this means that we are skipping data of 2.5% on the left side and 2.5% on the right side . now median of this data is 54.0000 . this data shows that the writing score mid point is 54.0000 this is the central point of the writing score. Now variance for the writing score which is square of standard deviation is 9.47859 . this variance shows that this is variation of writing score per unite. Now the minimum value that you have collected in this case is 31.00 which means that it is the lowest of the number in the data that is being collected and the maximum value is 67 which means that the highest number 67.00 is being recorded in this set of data. The range we are getting in this set of data is 36. If we see the range maximum minus minimum which becomes 36 but if we take inter quartile we need to do q1-q3 and through this we are getting the interquartile value that is 14.75. now the skewness value that is given is -.482 it is actually defining the shape. If the skewness is zero then ultimately distribution will

be normal and if the data is greater then 0 means that it is positively skewed and in this case it is less then 0 which is more like tends towards the negative skewed and the graph in this case will tends towards negative if we plot the graph. Kurtosis is .342 which means that it is the normal distribution.

 Q NO. 4:
  Explain The Following Normal Distribution?

The graph that s drawn is known as the normal distribution graph. On the left side all negative values lies and on the right side all the positive values lies.

Its symmetry means that 50% of cases lie to either side of the central point as defined by the mean. Two of the other most frequently-used representations are the portions lying between plus and minus one standard deviations of the mean (containing approximately 68% of cases,) and that between plus and minus 1.96 standard deviations (containing approximately 95% of cases,), sometimes rounded up to 2.00 for convenience. Thus, if a variable is normally distributed, we expect 95% of the cases to be within roughly 2 standard deviations from the mean. If we do +1 of the standard deviation then we will move ahead and if we do -1 then we will move behind if we go for 2 then 68% of the data will be covered and if we move -2 then we will cover the 95% of the data if we make it 3 standard deviation the 100% data will be covered.

Q NO.3 (PART A)
Interpret the following table

Now doing the interpretation for the given table of GENDER:

This data is being divided into two parts that MALE AND FEMALE and the total number of are 123 from which 62 are males and 61 are females. each of there frequency is being given.

Interpretation for the table for male and female :

The frequency for the males are give 62 out of 123 and the percent is 50.4 this shows that it has the higher percentage out of 100 so the graph that will be drawn greater . the valid percent is also the same 50.4 and the cumulative

Frequency is also 50.4. now the total numbers of the females are 61 and its percentage is 49.6 so the graph will going to be drawn shorter and it valid percent is also the 49.6 because there is no missing values. you add the cumulative percentage of males and females it make 100 .

So in this table the frequency of male are 62 which is the highest frequency and it percentage is 50.6 and it cumulative is also greater then females so the graph will move upward then females graph.

Q NO.3 (PART B)

IN this case as it is shown that the percentage for the male are 55.96 and the female percentage is 49.16 which is very much quite clear that the graph of the male is moving upwards then female just because of there percentages are show