

Assignment 1: Pay Equity

Objectives: At the completion of this assignment, you will be able to:

1. Apply advanced functionality and statistical tools of Microsoft Excel (v-lookup, pivot tables, etc) to identify potential, systemic gender bias in the company's compensation practices.

Instructions:

For this assignment, only one excel file called "Job Data Exercise," is needed. It includes several worksheets of data reflecting compensation data from your company.

Per VII of the CRA of 1964, the Equal Pay Act, and most recently the Lilly Ledbetter Act, employers are expected to conduct analyses of pay practices and identify potential sources of systemic bias against protected groups and institute reforms to remedy these biases. For the purpose of this study we will focus on gender so that we keep the complexity of the analysis to just two groups rather than multiple groups.

Steps include:

1. Using the VLOOKUP function, combine information contained on the DEMOGRAPHICS, JOB DATA, and SALARY INFO worksheets from the "Job Data Exercise" file, such that all data points per individual are included on a single worksheet.

Be careful to ensure that you convert data so that it is uniform for all incumbents (i.e., salary)... remember, there are 2080 hours in a standard work year, so to calculate an hourly rate from salary, you would divide the salary by the number of hours in the year. To convert from hourly to salary, you would do the opposite and multiply the hourly rate by the number hours in a year.

2. For the five jobs listed below, conduct a t-test (using the T.TEST function) for independent samples to determine if a statistically significant difference exists between the average compensation for men and women.

****DIFFERENT FROM THE WALKTHROUGH****

- CAD TECHNICIAN I
- ASSOCIATE CONSTRUCTION MANAGER
- ENGINEER I
- Sr. FIELD REP I
- PLANNER I

3. Draft a 500 word discussion of your findings. Include:
 1. A table indicating the avg hourly rate for men and women for each job (you will have to calculate these in EXCEL) and indicate if they are statistically different based on the t-test.
 2. Assume one or more **do** show significant differences to the detriment of women, what would you do to address the situation. Based on internet research, what is considered a best practice for correcting pay inequities while minimizing liabilities for the company?

Deliverables: To earn full credit, you will need to post the following in the assignment drop box.

1. Attach your paper to the assignment dropbox prior to the due date.
2. Attach the excel spreadsheet showing all your work including the computation of the t-test.

From the EXCEL HELP SCREEN:

VLOOKUP function

Searches for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify in the table. Use **VLOOKUP** instead of **HLOOKUP** when your comparison values are located in a column to the left of the data you want to find.

The V in **VLOOKUP** stands for "Vertical."

Syntax

VLOOKUP(lookup_value,table_array,col_index_num,range_lookup)

Argument	Description	Remarks
lookup_value	The value to be found in the first column of the array. The lookup_value argument can be a value, a reference, or a text string.	• If VLOOKUP can't find lookup_value, and range_lookup is TRUE, it uses the largest value that is less than or equal to lookup_value. • If VLOOKUP can't find lookup_value, and range_lookup is FALSE, VLOOKUP returns the #N/A value. • If lookup_value is smaller than the smallest value in the first column of table_array, VLOOKUP returns the #N/A error value.
table_array	The table of information in which data is looked up. You can use a reference to a range or a range name.	• The values in the first column of table_array can be text, numbers, or logical values. • If range_lookup is TRUE, the values in the first column of table_array must be placed in ascending order: ..., -2, -1, 0, 1, 2, ..., A-Z, FALSE, TRUE; otherwise VLOOKUP may not give the correct value. If range_lookup is FALSE, table_array does not have to be sorted. • You can put the values in ascending order by choosing the Sort command from the Data menu and selecting Ascending. • Uppercase and lowercase text are equivalent.
col_index_num	The column number in table_array from which the matching value must be returned.	• A col_index_num of 1 returns the value in the first column in table_array; a col_index_num of 2 returns the value in the second column in table_array, and so on. • If col_index_num is less than 1, VLOOKUP returns the #VALUE! error value. • If col_index_num is greater than the number of columns in table_array, VLOOKUP returns the #REF! error value.
range_lookup	A logical value that specifies whether you want this function to find an exact match or an approximate match.	• If TRUE or omitted, an approximate match is returned. In other words, if an exact match is not found, the next largest value that is less than lookup_value is returned. • If FALSE, this function will find an exact match. If one is not found, the error value #N/A is returned.

A GREAT (HR-RELEVANT) TUTORIAL ON THE VLOOKUP FUNCTION:

<https://www.youtube.com/watch?v=-BtsZ3CDIzo>

For this assignment:

- Lookup_value will be the ID of each employee
- Table_array will be the full range of the data on each of the worksheets you want to get data from
- Col_index_num will have to be adjusted for each column you want to pull into the summary sheet
- Range_lookup will always be **FALSE**

T.TEST function

Returns the probability that is associated with a Student's t-Test. Use **T.TEST** to determine whether two samples are likely to have come from the same two underlying populations that have the same mean.

Syntax

T.TEST(array1,array2,tails,type)

Argument	Description	Remarks
array1	The first data set.	None.
array2	The second data set.	None.
tails	Specifies the number of distribution tails.	- If tails = 1, T.TEST uses the one-tailed distribution. If tails = 2, T.TEST uses the two-tailed distribution. - If tails is any value other than 1 or 2, this function returns the #NUM! error value. - If this argument is nonnumeric, this function returns the #VALUE! error value. - If this argument contains a decimal value, this function ignores the numbers to the right side of the decimal point.
type	The kind of t-Test to perform.	- If type equals 1, T.TEST performs a paired test. If type equals 2, T.TEST performs a two-sample equal variance (homoscedastic) test. If type equals 3, T.TEST performs a two-sample unequal variance (heteroscedastic) test. - If this argument is nonnumeric, this function returns the #VALUE! error value. - If this argument contains a decimal value, this function ignores the numbers to the right side of the decimal point.

General remarks

- If **array1** and **array2** have a different number of data points, and **type** = 1 (paired), **T.TEST** returns the #N/A error value.
- T.TEST** uses the data in **array1** and **array2** to compute a non-negative t-statistic. If **tails**=1, **T.TEST** returns the probability of a higher value of the t-statistic under the assumption that **array1** and **array2** are samples from populations with the same mean. The value returned by **T.TEST** when **tails**=2 is double that returned when **tails**=1 and corresponds to the probability of a higher absolute value of the t-statistic under the "same population means" assumption.

FOR THIS ASSIGNMENT:

- Array 1 will be the wages for Females in the position
- Array 2 will be the wages for Males in the position
- Tails should **always be entered as 2** (we are doing a two-tailed test of statistical significance)
- Type should **always be entered as 2** (we are assuming that there is equal variance in the female wages and the male wages)