

**INFO 6540 Data Management (Winter 2019)**  
**Assignment 1: Unstructured to Structured Data**

**Submitted by,**  
**Vijaya Lakshmi Esakkimuthu**  
**B00813377**  
**Master's in Resource and Environmental Management**  
**Dalhousie University**

**For Professor. Elvira Mitraka**  
**School of Information Management**  
**Dalhousie University**

**Answers:**

**1. Find which actors have more than one Academy Award in the categories: Academy Award for Best Actress and Academy Award for Best Actor, and how many they have?**

In this question the number of awards for the best actors and actress was found using the COUNTIFS formula. Concatenate function was used to display the name of the actor/actress and the number of awards. For each year the best actor/actress was found using the Index Match function. The formula for index match is found in cell C4 and E4 and the formula for COUNTIFS is found in cells D4 and F4.

**2. For each year, calculate the number of male and female winners in total (all categories).**

The number of male and female award winners was calculated for each year from the given data set.

- COUNTIFS formula was used to get the number of male and female winners in each year. The COUNTIFS formula could be used for multiple ranges and multiple criteria.
- The range and criteria for this formula was chosen from the main dataset and the number of male winners was compared with each year. Ex: = COUNTIFS (range of cells from gender, male, year, each year).
- The formula for this question could be found in cell D4 for the number of males in each year and cell E4 for number of female winners in each year.

**3. Calculate how many female and male winners there are overall.**

The total number of male and female in the dataset was found each separately using the COUNTIF formula and the overall sum was calculated.

- The formula for this could be found in cells C3, C4, and the formula for total sum is found in C6.

**4. Find which directors have more than one Academy Award in the category: Academy Award for Best Director, and how many they have**

In this question the best director award for each year was found using the Index Match function. The formula for this is found in cell C5. The number of awards was found using COUNTIFS function and the formula could be found in cell D5.

**5. Find how many female directors won an Academy Award in the category: Academy Award for Best Director.**

To find the number of female directors who won the Academy Award for best Director, the COUNTIFS formula was used. The Gender column was selected to get the range of cells from B2 to B1862 and was compared with gender Female in cell reference B5 in sheet 5. The second criteria used in the COUNTIFS formula was Category from cells D2 to D1862 from oscarsdata sheet and cell D5 from Ans 5 sheet.

**6. Find which movie won the most Academy Awards (all categories included). If there is a tie, list the titles of the movies in alphabetical order.**

In order to find the movies which won the most number of Academy awards, the number of awards won by each movie was calculated. The CountIF function was used to find the number of number of times of the movie won the award. The list of movie names from the oscarsdata.csv was copied to the new sheet6. The Index label was used to find the movie with the maximum number of awards. The formula for the number of awards for each movie is found in cell B2 and the formula for finding the maximum number of movies from Column A 2: A 1861 and is matched with column B2: B186.

**7. Make a list with movies that have been awarded only one Academy Award.**

For this the data from answer 6 was used. The IF condition was used to check the movies with only one award. The formula is found in cell E6.