**Last Name, Name: …………………………………………………………………………………………………………….**

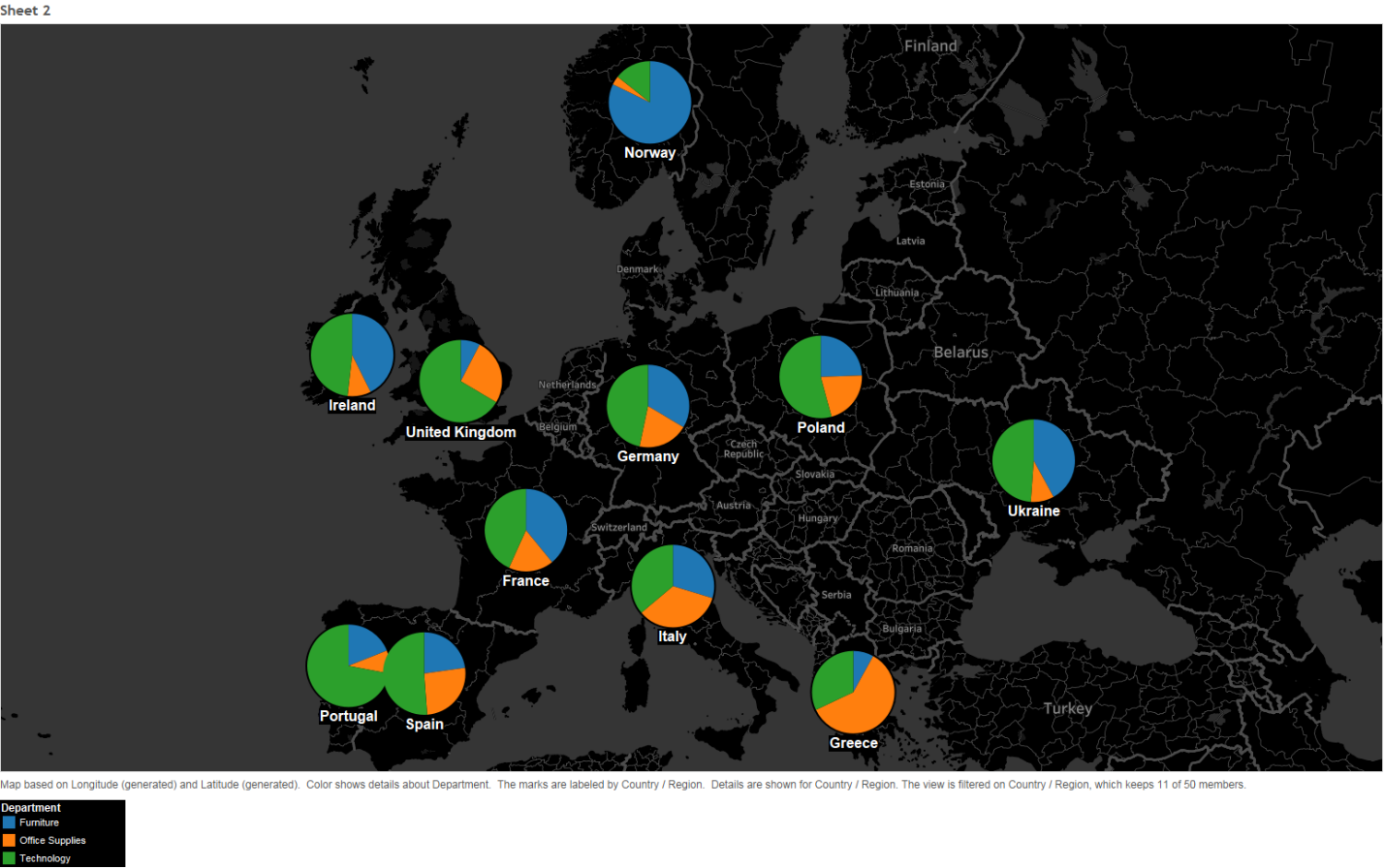
The data needed for Problems 1-3 is the **Heat Map starter of chapter 9** of the fundamental workbook.

Please be reminded of Academic Integrity Policy of WSOM in which there is zero tolerance for any kind of cheating.

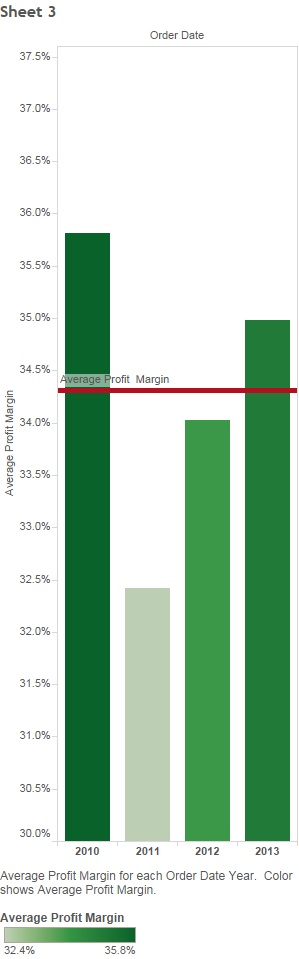
Problem 1 (20 Points) (**Output In:** Tableau, Create Worksheet named “1”): Create the following visualization in Tableau.



Problem 2 (10 Points) (**Output In:** Tableau; Create Worksheet named “2”): Create the following visualization in Tableau.



Problem 3 (10 Points) (**Output In:** Tableau; Create Worksheet named “3”): Create the following visualization in Tableau.



Problem 4 (20 Points) (**Output In:** Tableau; Create Worksheet named “4”)

**Data:** “2014\_Olympics.xlsx” available in Tableau data sets used in teaching Tableau in class

**Tasks:**

Build a vertical side by side bar chart comparing the percentage of total medals of Freestyle (first bars) vs Alpine Skiing (second bars) won by athletes in four regions: North America (Canada and the US), Australia, Asia (China, Kazakhstan, South Korea, Japan), and Europe (all else). Slicing to be done first on region and then by sport (Freestyle and Alpine) yielding side by side bars of Freestyle and Alpine per region (if available). The percentage value of a region for a sport should represent the percentage of total medals of that region won in that sport. These percentages should be shown with one decimal on the bars. Color the bars by sport type. Properly scale and name the y-axis (shown with one decimal). Properly name the chart.

Problem 5 (20 Points) (**Output In:** Tableau; Create Worksheet named “5”)

**Data:** “2014\_Olympics.xlsx” available in Tableau data sets used in teaching Tableau in class

**Tasks:**

Build a tree map of total gold medals won sliced by country (first) and sport (second). Provide labels properly on each slice.

Problem 6) (20 Points) (**Output In:** R-Studio Only)

**Data:** “Airfares.csv” available in Book Data set (CANVAS /Book Dataset.zip)

**Tasks:** Write codes that

1. build and display a new data frame that calculates average, median, and standard deviation of Fare values.
2. Partition the data into 70% training and 30% validation. Use training data to build a linear regression model to predict FARE based on DISTANCE. Compute MAPE on validation data using this model.