

**Economics 30330: Statistics for Economics**  
**Problem Set 1**  
*University of Notre Dame*  
**Instructor: Julio Garín**  
**Spring 2012**

**Due Date:** Beginning of class on Wednesday, February 1st. Please complete the assignment in the allotted space. You may work in groups, but you need to turn your own work. You may also draw all the figures by hand.

**Introduction to Data and Statistics (60 points)**

1. Part of the data bank of a corporation is shown below.

<b>Employee Number</b>	<b>Gender</b>	<b>Department</b>	<b>Years of Experience</b>	<b>Employee rank (1-10)</b>	<b>Yearly Salary</b>
23450	Male	Accounting	15	10	\$ 52,443.00
34568	Female	IT	24	7	\$ 111,239.00
23123	Female	Personnel	20	4	\$ 84,473.00
23007	Male	Finance	9	1	\$ 47,519.00

For each of the following, indicate whether the variable is an example of nominal, ordinal, interval, or ratio scale data, and explain your answer:

(a) ‘Employee Number’

(b) ‘Gender’

(c) ‘Years of Experience’

(d) ‘Employee Rank’

(e) ‘Yearly Salary’

2. The following data shows the yearly income distribution of a sample of 200 employees at MNM, Inc.

<b>Yearly Income (Thousands)</b>	<b>Number of Employees</b>
20-24	2
25-29	48
30-34	60
35-39	80
40-44	10

- (a) How many variables are presented in the above data set?
- (b) The above data set represents the results of how many observations?
- (c) What percentage of employees has yearly incomes of \$35,000 or more?
- (d) Is the figure (percentage) that you computed in Part *c* an example of statistical inference? If no, what kind of statistics does it represent? Justify.
- (e) Based on this sample, the president of the company said that “45% of all our employees’ yearly incomes are \$35,000 or more.” The president’s statement represents what kind of statistics? Justify.
- (f) With the statement made in Part *c*, can we be assured that more than 45% of all employees’ yearly incomes are at least \$35,000? Explain.
- (g) What percentage of employees of the sample has yearly incomes of \$29,000 or less?

3. The following shows the temperatures (high, low) and weather conditions on a given Sunday for some selected world cities. For the weather conditions, the following notations are used: *c* = clear; *cl* = cloudy; *sh* = showers; *pc* = partly cloudy.

<b>City</b>	<b>High</b>	<b>Low</b>	<b>Condition</b>
Acapulco	99	77	<i>pc</i>
Bangkok	92	78	<i>pc</i>
Mexico City	77	57	<i>sh</i>
Montreal	72	56	<i>pc</i>
Paris	77	58	<i>c</i>
Rome	88	68	<i>cl</i>
Toronto	78	61	<i>c</i>

- (a) How many elements are in this data set?
- (b) How many variables are in this data set?
- (c) How many observations are in this data set?

- (d) Name the variables and indicate whether they are qualitative or quantitative.
- (e) For which variables are arithmetic operations appropriate and for which are they not appropriate?

4. The frequency distribution below was constructed from data collected from a group of 25 students.

Height (In inches)	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
58-63	3			
64-69	5			
70-75	2			
76-81	6			
82-87	4			
88-93	3			
94-99	2			

- (a) Complete the missing columns.
- (b) Plot the Relative Frequency Histogram and the Cumulative Relative Frequency Histogram.

5. The following data set shows the number of hours of sick leave that some of the employees of Lay Z Inc. have taken during the first quarter of the year (rounded to the nearest hour).

19	22	27	24	28	12
23	47	11	55	25	42
36	25	34	16	45	49
12	20	28	29	21	10
59	39	48	32	40	31

- (a) Develop a frequency distribution, a relative frequency distribution, a percent frequency distribution, and a cumulative frequency for the above data. Let the width of your classes be 10 units and start your first class as 10 - 19.

<b>Hours Sick Leave Taken</b>	<b>Frequency</b>	<b>Relative Frequency</b>	<b>Percent Frequency</b>	<b>Cumulative Frequency</b>
10-19				

(b) Plot the Cumulative Frequency.

6. A survey of 400 college seniors resulted in the following crosstabulation regarding their undergraduate major and whether or not they plan to go to graduate school.

<b>Graduate School</b>	<b>Business</b>	<b>Engineering</b>	<b>Others</b>	<b>Total</b>
Yes	35	42	63	140
No	91	104	65	260
Total	126	146	128	400

(a) Are a majority of the seniors in the survey planning to attend graduate school? Explain.

(b) Which discipline constitutes the majority of the individuals in the survey? Explain.

(c) Compute row percentages and comment on the relationship between the students' undergraduate major and their intention to attend graduate school.

<b>Graduate School</b>	<b>Business</b>	<b>Engineering</b>	<b>Others</b>	<b>Total</b>
Yes				
No				
Total				

(d) Compute the column percentages and comment on the relationship between the students' intention to go to graduate school and their undergraduate major.

<b>Graduate School</b>	<b>Business</b>	<b>Engineering</b>	<b>Others</b>	<b>Total</b>
Yes				
No				
Total				