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

Introduction to Data and Statistics (60 points)

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

Employee Number	Gender	Department	Years of Experience	Employee rank (1-10)	Yearly Salary
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'
Nominal. It refers to a categorical data in which the order of the categories is arbitrary.

(b) 'Gender'
Nominal. Same reason as before; in this case it is even more obvious since 'gender' is just a label.

(c) 'Years of Experience'
Ratio. Years of experiences refers to a continuous data and both the difference and the ratio between two years have clear interpretation.

(d) 'Employee Rank'
Ordinal. In this case, it is a categorical data in which there is a meaningful order.

(e) 'Yearly Salary'
Ratio. Same as in part c.

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

Yearly Income (Thousands)	Number of Employees
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?
One: yearly income.

(b) The above data set represents the results of how many observations?
200 employees.

(c) What percentage of employees has yearly incomes of \$35,000 or more?
$\frac{80+10}{200}\% = 45\%$

(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.
Is not inferring from the sample, it is just describing it. Therefore, it is descriptive statistics.

(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.
In this case, she is clearly inferring. Hence it represents a case of statistical inference.

(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.
No, this is simply an inference and approximation based on the sample information.

(g) What percentage of employees of the sample has yearly incomes of \$29,000 or less?
$\frac{48+2}{200}\% = 25\%$

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.

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

(a) How many elements are in this data set?
Seven: each city is an element.

(b) How many variables are in this data set?
Three: Hi, Low, and Condition.

(c) How many observations are in this data set?
Seven: for each city there is one observation.

(d) Name the variables and indicate whether they are qualitative or quantitative.
Hi: quantitative, Lo: quantitative, Condition: qualitative.

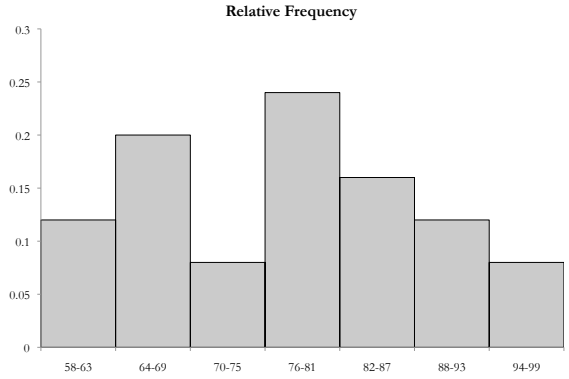
(e) For which variables are arithmetic operations appropriate and for which are they not appropriate?
Hi: appropriate, Lo: appropriate, Condition: not appropriate.

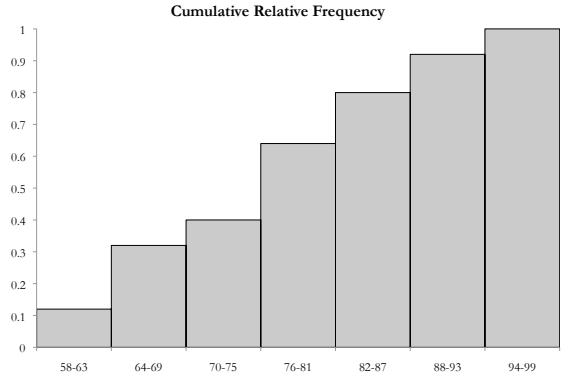
4. The frequency distribution below was constructed from data collected from a group of 25 students.

Height (In inches)	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
58-63	3	0.12	3	0.12
64-69	5	0.20	8	0.32
70-75	2	0.08	10	0.40
76-81	6	0.24	16	0.64
82-87	4	0.16	20	0.80
88-93	3	0.12	23	0.92
94-99	2	0.08	25	1.00

(a) Complete the missing columns.

(b) Plot the Relative Frequency Histogram and the Cumulative Relative Frequency Histogram.





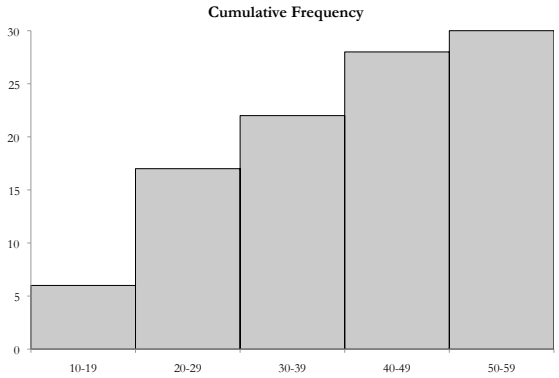
5. The following data set shows the number of hours of sick leave that some of the employees of Lay Z Inc. have taken during the first quarter of the year (rounded to the nearest hour).

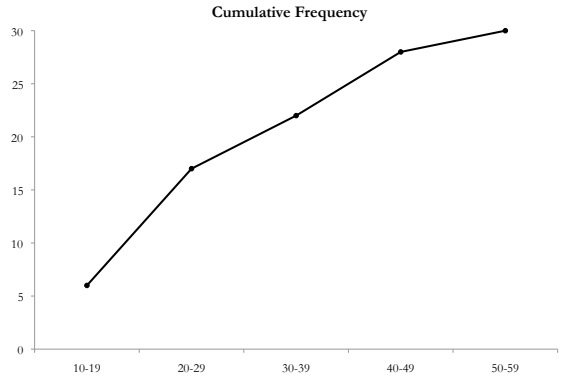
19	22	27	24	28	12
23	47	11	55	25	42
36	25	34	16	45	49
12	20	28	29	21	10
59	39	48	32	40	31

(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.

Hours Sick Leave Taken	Frequency	Relative Frequency	Percent Frequency	Cumulative Frequency
10-19	6	0.20	20%	6
20-29	11	0.37	37%	17
30-39	5	0.17	16%	22
40-49	6	0.20	20%	28
50-59	2	0.07	7%	30

(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.

Graduate School	Business	Engineering	Others	Total
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.
No, majority (260) will not attend graduate school.

(b) Which discipline constitutes the majority of the individuals in the survey? Explain.
The majority (146) are engineering majors.

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

Graduate School	Business	Engineering	Others	Total
Yes	25%	30%	45%	100%
No	35%	40%	25%	100%
Total	31.5%	36.5%	32%	100%

The majority who plan to go to graduate school are from 'Other' majors; majority of those who will not go to graduate school are Engineering majors.

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

Graduate School	Business	Engineering	Others	Total
Yes	27.8%	28.8%	49.2%	35%
No	72.2%	71.2%	50.8%	65%
Total	100%	100%	100%	100%

Approximately the same percentages of Business and Engineering majors plan to attend graduate school (27.8% and 28.8% respectively). Of the 'Other' majors approximately half (49.2%) plan to go to graduate school.