



commerce
undergraduate
society

COMM 205 MANAGEMENT INFO SYSTEMS SOLUTION PACKAGE

BY LEAH ZHANG



SAMPLE Q'S ANSWER KEY

Use the following scenario and Stata screenshot to answer Questions 1 to 4.

The dataset air.dta with 10 observations containing airports' names, year, number of passengers (in millions) is displayed below. It shows 5 airports in Canada, and their number of passengers in a year for 2014 and 2015. The variable airportsname is a string variable, while the variables year and numberofpassengersinmillions are numerical variables.

	airportsname	year	numberofpa~s
1	YXX	2014	.4
2	YAA	2014	1.1
3	YVR	2014	1.8
4	YYJ	2014	.7
5	YGK	2014	.9
6	YXX	2015	.6
7	YAA	2015	.9
8	YVR	2015	1.7
9	YYJ	2015	1.2
10	YGK	2015	1.3

- Right after browsing the data above, Jessica ran a Stata command, and got the following output:

number of passengers in millions				
Percentiles		Smallest		
1%	1.7	1.7		
5%	1.7	1.8		
10%	1.7	.	Obs	2
25%	1.7	.	Sum of Wgt.	2
50%	1.75		Mean	1.75
		Largest	Std. Dev.	.0707107
75%	1.8	.		
90%	1.8	.	Variance	.005
95%	1.8	1.7	Skewness	0
99%	1.8	1.8	Kurtosis	1

Which one of the following formulas, when properly executed, will result in the above output?

- sum numberofpassengersinmillions if airportsname=="YVR" & (year=2014 | year=2015)
- sum numberofpassengersinmillions if airportsname=="YVR", detail
- sum numberofpassengersinmillions if (airportsname=="YVR" & year==2014) | year ==2015, detail
- sum numberofpassengersinmillions if year==2014 | (year==2015 & airportsname=="YVR"), detail



2. Then you ran the following command in Stata:

```
tab airportsname, sum (numberofpassengersinmillions) if year==2014
```

Stata returned an error message when the above command was running. Jessica came to you for help. You fixed the command, reran it on Stata, and got the following output:

airports name	Summary of number of passengers in millions		
	Mean	Std. Dev.	Freq.
YAA	1.1	0	1
YGK	.9	0	1
YVR	1.8	0	1
YXX	.4	0	1
YYJ	.7	0	1
Total	.98	.52630789	5

In 20 words or fewer, explain on the answer sheet how you would fix the above erroneous formula. You do NOT need to write the correct command. If you prefer, however, you may write the correct command instead.

Move the if year==2014 between the word airportsname and the comma.

```
tab airportsname if year==2014, sum (numberofpassengersinmillions)
```



3. Jessica sorted the data using one Stata command. After she ran the Stata command to sort the dataset, she ran the browse command and got the following output:

	airportsname	year	numberofpa~s
1	YVR	2015	1.7
2	YGK	2015	1.3
3	YYJ	2015	1.2
4	YAA	2015	.9
5	YXX	2015	.6
6	YVR	2014	1.8
7	YAA	2014	1.1
8	YGK	2014	.9
9	YYJ	2014	.7
10	YXX	2014	.4

Write a single Stata command that, when properly executed, will sort the dataset in the above order.

Answer:

```
gsort -year -numberofpassengersinmillions
```



4. Suppose that whatever data Jessica had opened before is now cleared, and she reopened air.dta. **In 20 words or fewer, explain what would happen if Jessica then ran the following command:**

```
drop if airportsname="YVR"
```

Stata will return an error message.



5. The following screenshot includes 5 commerce courses (subject): COMM205, 298, 370, 204, and 101. Every year (year), the system records the number of students that are enrolled in these courses (enrolled), and the number of students who pass the courses (passed). The following Stata screenshot provides you with the information in 2014 and 2015. There are only 10 observations in the following Stata screenshot, in a dataset called courses.dta. Note that the variable subject is a string variable; the other 3 variables are numerical variables.

	subject	year	enrolled	passed
1	COMM205	2014	501	501
2	COMM298	2014	.	100
3	COMM370	2014	376	360
4	COMM204	2014	58	40
5	COMM101	2014	90	89
6	COMM205	2015	623	623
7	COMM298	2015	323	300
8	COMM370	2015	498	478
9	COMM204	2015	33	32
10	COMM101	2015	189	.

You are asked to create a new variable called *incr_ps_2014*, which should be equal to 1 if the number of passed students (passed) in that subject increased from 2014 to 2015, and 0 otherwise. If the number of passed students in that subject decreased or did not change between 2014 and 2015, then *incr_ps_2014* should be equal to 0.

Furthermore, *incr_ps_2014* should be missing for all observations where year is equal to 2014. *incr_ps_2014* should only be either 0 or 1 where year is equal to 2015.

In the answer sheet, write at most 3 Stata commands that would allow you to complete the above task.

```
bysort subject: egen a = mean(passed)
```

```
gen incr_ps_2014 = 1 if year==2015 & passed>a
```

```
replace incr_ps_2014 = 0 if year==2015 & incr_ps_2014==.
```

