

Lab 17 *Some of these problems may be more challenging than others. Please feel free to work with others, attend office hours, or post on the course discussion forum if you need help. While collaboration with other students is encouraged, each student is responsible for submitting his or her own work. This assignment should be submitted in one well-commented SAS program. For any questions that require a written answer, do so in the SAS comments. Be sure to re-name the uploaded SAS scripts according to the naming convention `LastnameFirstinitial_Lab#.sas` (e.g., `PileggiS_Lab17.sas`).*

Use the 2012 Olympic Medalists data set `02012.sas7bdat` for the following exercises (this data set name starts with an “oh” and not a zero). **Example output is provided on the last pages.**

1. Create a SAS library to access the `02012.sas7bdat`.
2. While for certain sports athletes peak at younger ages, in some sports there is more variability in the age of athletes. Using `PROC SQL`, create a summary of the olympics data by sport that
 - (a) displays the sport name
 - (b) displays the number of medalists that participated in each sport
 - (c) displays the standard deviation of athlete’s age for each sport
 - (d) is sorted by standard deviation in descending order

Which sport has the largest variability in age? Note your findings in a comment in your SAS code.

3. Suppose all Olympic medalists of the same country are traveling together and need to get on an elevator that has a 2000kg weight limit. Which countries could fit all of their medalists on this elevator? Using `PROC SQL`, create a summary of the olympics data by country that
 - (a) displays the country name
 - (b) displays the number of medalists for each country
 - (c) sums the total weight of all medalists per country
 - (d) creates a new variable called `elevator` which has a value of “yes” if the total weight is less than or equal to 2000, and “no” otherwise; make sure that countries with a missing value for total weight also have a missing value for `elevator`
 - (e) is sorted by total weight in descending order

In examining the output, how many countries wouldn’t be able to put all of their medalists on the elevator because of the weight limit? Note your findings in a comment in your SAS code.

4. Using `PROC SQL`, re-create the table from the previous two labs where you summarized the olympics data by country. Display the following variables:
- (a) country
 - (b) number of medalists in each country
 - (c) total number of medals won in each country
 - (d) average number of medals won per athlete in each country
 - (e) maximum number of medals won per athlete in each country
 - (f) average age of medalists in each country
 - (g) average weight of medalists in each country
 - (h) percent of medalists that are female in each country
 - (i) percent of medalists that are male in each country

Hint: For the last two variables regarding the percent of males and the percent of females, use a `CASE` statement within a `SUM` expression such as

```
SUM(CASE WHEN condition THEN value1 ELSE value2 END) AS newvariable
```

Question 2

Sport	N	sd_age
Equestrian	16	9.439059
Shooting	28	7.440416
Cycling - Road	8	5.655276
Canoe Slalom	15	4.730851
Rowing	117	4.637524
Fencing	24	4.247548
Badminton	5	4.159327
Diving	23	4.094728
Archery	22	3.859267
Swimming	97	3.815578
Cycling - Track	15	3.81351
Judo	53	3.611082
Gymnastics - Artistic	29	3.572955
Weightlifting	29	3.034287
Cycling - Road, Cycling - Track	2	2.828427
Table Tennis	6	2.065591
	3	1.527525
Trampoline	3	0
Cycling - Mountain Bike, Cycling - Track	1	.

Question 3

Country	N	sum_weight	elevator
United States of America	61	4373	no
Great Britain	45	3196	no
Germany	38	3025	no
People's Republic of China	50	2911	no
France	27	2078	no
Canada	25	1944	yes
Australia	25	1829	yes
Russian Federation	29	1741	yes
Republic of Korea	19	1335	yes
Japan	22	1215	yes
Netherlands	17	1191	yes
New Zealand	14	1106	yes
Italy	14	1069	yes
Denmark	7	537	yes
Ukraine	7	478	yes
South Africa	6	434	yes
Brazil	5	408	yes
Slovenia	4	324	yes
Mexico	6	319	yes
Romania	10	316	yes
Cuba	5	312	yes
Slovakia	4	288	yes
Democratic People's Republic of Korea	5	286	yes
Croatia	4	282	yes
Hungary	4	275	yes
Belarus	4	273	yes
Kazakhstan	4	259	yes
Poland	3	196	yes
Colombia	3	194	yes
Mongolia	2	173	yes
Czech Republic	2	171	yes
Sweden	2	140	yes
Belgium	2	122	yes
Spain	3	115	yes
Greece	1	90	yes
Serbia	1	85	yes
Egypt	1	82	yes
Venezuela	1	80	yes
Qatar	1	70	yes
Georgia	1	66	yes
Lithuania	1	64	yes
Indonesia	1	62	yes
Thailand	1	58	yes
Singapore	1	57	yes
Azerbaijan	1	56	yes
Republic of Moldova	1	53	yes

Question 4

Country	N	total_sum	total_mean	total_max	ave_age	ave_weight	pct_female	pct_male
Australia	25	32	1.28	4	24.4	73.2	76.0%	24.0%
Azerbaijan	1	1	1.00	1	18.0	56.0	0.0%	100%
Belarus	4	4	1.00	1	29.5	68.3	75.0%	25.0%
Belgium	2	2	1.00	1	26.0	61.0	50.0%	50.0%
Brazil	5	5	1.00	1	23.4	81.6	40.0%	60.0%
Canada	25	25	1.00	1	28.2	77.8	56.0%	44.0%
Colombia	3	3	1.00	1	26.7	64.7	33.3%	66.7%
Croatia	4	4	1.00	1	23.5	94.0	0.0%	100%
Cuba	5	5	1.00	1	26.4	78.0	40.0%	60.0%
Czech Republic	2	2	1.00	1	27.0	85.5	0.0%	100%
Democratic People's Republic of Korea	5	5	1.00	1	23.0	57.2	60.0%	40.0%
Denmark	7	7	1.00	1	29.1	76.7	14.3%	85.7%
Egypt	1	1	1.00	1	21.0	82.0	0.0%	100%
France	27	33	1.22	3	25.1	77.0	48.1%	51.9%
Georgia	1	1	1.00	1	20.0	66.0	0.0%	100%
Germany	38	40	1.05	2	27.5	81.8	28.9%	71.1%
Great Britain	45	44	1.00	1	28.4	81.9	24.4%	73.3%
Greece	1	1	1.00	1	25.0	90.0	0.0%	100%
Hungary	4	4	1.00	1	26.3	68.8	25.0%	75.0%
India	2	2	1.00	1	27.5	.	0.0%	100%
Indonesia	1	1	1.00	1	23.0	62.0	0.0%	100%
Italy	14	17	1.21	2	29.4	76.4	35.7%	64.3%
Japan	22	24	1.09	2	23.6	67.5	40.9%	59.1%
Kazakhstan	4	4	1.00	1	27.3	64.8	75.0%	25.0%
Lithuania	1	1	1.00	1	15.0	64.0	100%	0.0%
Mexico	6	6	1.00	1	20.2	53.2	66.7%	33.3%
Mongolia	2	2	1.00	1	27.0	86.5	0.0%	100%
Netherlands	17	18	1.06	2	27.4	70.1	94.1%	5.9%
New Zealand	14	14	1.00	1	32.5	79.0	28.6%	71.4%
Norway	2	2	1.00	1	25.0	.	0.0%	100%
Pakistan	1	.	.	.	24.0	.	0.0%	0.0%
People's Republic of China	50	53	1.06	3	23.0	66.2	48.0%	52.0%
Poland	3	3	1.00	1	27.3	65.3	100%	0.0%
Qatar	1	1	1.00	1	41.0	70.0	0.0%	100%
Republic of Korea	19	22	1.16	2	26.8	70.3	47.4%	52.6%
Republic of Moldova	1	1	1.00	1	19.0	53.0	100%	0.0%
Romania	10	10	1.00	1	22.3	63.2	80.0%	20.0%
Russian Federation	29	32	1.10	2	23.7	75.7	44.8%	55.2%
Serbia	1	1	1.00	1	34.0	85.0	0.0%	100%
Singapore	1	1	1.00	1	25.0	57.0	100%	0.0%
Slovakia	4	4	1.00	1	31.0	72.0	25.0%	75.0%
Slovenia	4	4	1.00	1	38.0	81.0	25.0%	75.0%
South Africa	6	6	1.00	1	24.5	72.3	0.0%	100%
Spain	3	2	1.00	1	25.3	57.5	66.7%	0.0%
Sweden	2	2	1.00	1	41.5	70.0	50.0%	50.0%
Taipei (Chinese Taipei)	1	1	1.00	1	21.0	52.0	100%	0.0%