

Homework 4

Due: Thursday, October 21st (before 11:59 pm) on GauchoSpace.

Instructions: For the General Knowledge Questions, answer as succinctly as possible. For the Programming Assignment, show a partial screenshot of your output. Submit your code separately as a SAS program file.

General Knowledge Questions

- the general format is: Variable_(name) = expression;
1. How do you create a new variable? Use variable assigning statements in the DATA step.
 2. Give an example of a *formatted* input statement. input @1 date\$ mmddyy. @5 name\$ 4.
 3. What are two symbols used for absolute pointer control? @n and #n
 4. In which step(s) are both WHERE and IF statements valid? In a DATA set
 5. What is the SAS date value for October 21, 2021? 22574
 6. The DATA step is processed in how many phases? List the phases. two phases - compile and execution
 7. Which SAS function returns the day of the week of a SAS date value? weekday(sas-date)
 8. What is a SAS date constant? It returns a SAS date value for a specific date
 9. By default, how many executable statements are allowed per IF-THEN or ELSE statement? only one executable statement
 10. Which window can be used to confirm the number of records and variables being read from and written to the SAS data set(s)? log window

Programming Assignment

1. Use the **avengers.xlsx** file posted on GauchoSpace.
 - a) Write an **IMPORT** procedure to create a SAS data set named **avengers** from the Excel file. Import the worksheet **Characters**.
 - b) Create a listing report using the **avengers** data set. Display the date and time, and display page numbers starting at 10.
 - c) Now create an html report displaying the descriptor portion of the **avengers** data set.
2. Use the **avengers** data set created in Exercise 1. This data set contains information about Avenger characters over the years.
 - a) Create a temporary SAS data set named **avengers2** by duplicating the **avengers** data set. Then do the following in **avengers2**:
 - i. Create a variable named **YearIntro** that extracts the year from the variable **AvengersIntro**. If the value for **AvengersIntro** is missing, set value of **YearIntro** to be 1963.

- ii. Create a variable named **Days** that contains the number of days since the character was introduced to October 1st, 2021. If the value for **AvengersIntro** is missing, assume a value of September 1st, 1963.
 - iii. Create a variable named **Years** that contains the number of years since the character was introduced to October 1st, 2021 (assume one year = 365.25 days). Hint: **floor(x)** rounds **x** down to the nearest integer.
 - iv. Create a variable named **Gender** that equals 'Male' when **Sex**=0 and 'Female' when **Sex**=1.
 - v. The **avengers2** dataset should contain only the variables **Name**, **Appearances**, **Gender**, **AvengersIntro**, **YearIntro**, **Days**, and **Years**.
- b) Now create a report using the **avengers2** data set
- i. Display the variables **Name**, **Gender**, **Appearances**, **AvengersIntro**, **YearIntro**, and **Years** (in this order).
 - ii. Display the values of **Appearances** with commas.
 - iii. Display the values of **AvengersIntro** in the format of 10/21/21.
 - iv. Add an appropriate title.
 - v. Display 'Character Name' as the column header for **Name**, 'Number of Appearances' for **Appearances**, 'Date Introduced' for **AvengersIntro**, 'Year Introduced' for **YearIntro**, and 'Years Since Introduction' for **Years**.
- c) Look at the report from part B. What is the problem with the values of Gender? *Female gets cut off.*
- d) Modify the code from part A to fix this problem (highlight the updated code). Hint: *cut off.* use a format. Rerun your updated code and the code from part B. Check the output to see if the problem has been fixed. *Format Gender \$6.; - added to 2a code!*
- e) Use the updated **avengers2** data set from part D.
- i. Create a SAS data set named **CharPopularity** from **avengers2** that only reads in the variables **Name**, **AvengersIntro**, and **Appearances**.
 - ii. Create a variable named **Popularity** that contains the popularity for each character. Assign the **Popularity** values as follows:
 - 'Low' when the appearances is less than 100
 - 'Medium' when the appearances is at least 100 and less than 1000.
 - 'High' when the appearances is at least 1000

*Note: Make sure to format the variable **Popularity** appropriately so that each value is fully displayed.*
 - iii. Permanently assign a format to **AvengersIntro** that displays the values in the format of 21OCT2021, and a format to **Appearances** that displays the values with commas.
- f) Now create a report using the **CharPopularity** data set.

- Group the values by **Popularity**. Display the different groups on separate pages.
- Set the **Popularity** as an identification variable.
- Only display **Name** and **AvengersIntro**.
- Add an appropriate title.
- Display 'Character Name' as the column header for **Name**, and 'Date Introduced' for **AvengersIntro**.
- Display the number of observations for each group at the bottom of each group.
- In what order are the **Popularity** groups displayed? Why?

Displays from high to low because ascending order is defaulted. write 'decending' after by statement before variable.

1a)

Obs	Name	Appearances	Current	Sex	ProbationaryIntro	AvengersIntro	Honorary	Death1	Return1	Death2	Return2	Death3	Return3	Death4	Return4	Death5	Return5
1	Henry Jonathan "Hank" Pym	22000	YES	0	.	SEP1963	Full	YES	NO								
2	Janet van Dyne	1165	YES	1	.	SEP1963	Full	YES	YES								
3	Anthony Edward "Tony" Stark	3068	YES	0	.	SEP1963	Full	YES	YES								
4	Robert Bruce Banner	2089	YES	0	.	SEP1963	Full	YES	YES								
5	Thor Odinson	2402	YES	0	.	SEP1963	Full	YES	YES	YES	NO						
6	Richard Milhouse Jones	612	YES	0	.	SEP1963	Honorary	NO									
7	Steven Rogers	3458	YES	0	.	MAR1964	Full	YES	YES								
8	Clinton Francis Barton	1456	YES	0	.	MAY1965	Full	YES	YES	YES	YES						
9	Pietro Maximoff	769	YES	0	.	MAY1965	Full	YES	YES								
10	Wanda Maximoff	1214	YES	1	.	MAY1965	Full	YES	YES								
11	Jacques Duquesne	115	NO	0	.	SEP1965	Full	YES	YES								
12	Heraclius	741	YES	0	.	OCT1967	Full	NO									
13	T'Challa	780	NO	0	.	MAY1968	Full	NO									
14	Victor Shade (alias)	1036	YES	0	.	NOV1968	Full	YES	YES								
15	Dane Whitman	482	NO	0	.	DEC1969	Full	NO									
16	Natalia Alianovna Romanova	1112	YES	1	.	MAY1973	Full	YES	YES								
17	Brandt	160	NO	1	.	AUG1973	Full	YES	YES								
18	Henry P. McCoy	1886	NO	0	JUL1975	SEP1976	Full	NO									
19	Heather Douglas	332	NO	1	JUL1975	SEP1976	Full	YES	YES	YES	YES						
20	Patsy Walker	557	NO	1	NOV1975	SEP1976	Full	YES	YES								
21	Matthew Liebowitz (birth name)	197	NO	0	.	AUG1975	Full	YES	YES	YES	NO						
22	Robert L. Frank Sr.	106	NO	0	APR1977	JUL1978	Full	YES	NO								
23	Simon Williams	692	YES	0	APR1977	MAR1979	Full	YES	YES								
24	Yondu Udonta	109	NO	0	.	FEB1978	Honorary	NO									
25	Marlinex T'Naga	100	NO	0	.	FEB1978	Honorary	NO									
26	Charlie-27	132	NO	0	.	FEB1978	Honorary	NO									
27	Nichollette Gold	108	NO	1	.	FEB1978	Honorary	NO									
28	Stakar	100	NO	0	.	FEB1978	Honorary	NO									
29	Vance Astrovik	156	NO	0	.	FEB1978	Honorary	NO									
30	Mar-Vell	254	NO	0	.	JUL1978	Full	YES	YES	YES	YES	YES	NO				
31	Carol Susan Jane	935	YES	1	.	APR1979	Full	NO									

The CONTENTS Procedure

Data Set Name	WORK.AVENGERS	Observations	163
Member Type	DATA	Variables	17
Engine	V9	Indexes	0
Created	10/21/2021 14:25:45	Observation Length	112
Last Modified	10/21/2021 14:25:45	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information

Data Set Page Size	131072
Number of Data Set Pages	1
First Data Page	1
Max Obs per Page	1168
Obs in First Data Page	163
Number of Data Set Repairs	0
Filename	/saswork/SAS_work20EC00006D15_0daws01-usw2.oda.sas.com/SAS_work15DB00006D15_0daws01-usw2.oda.sas.com/avengers.sas7bdat
Release Created	9.0401M6
Host Created	Linux
Inode Number	1074827678
Access Permission	rw-r--r--
Owner Name	u59557716
File Size	256KB
File Size (bytes)	262144

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat	Label
2	Appearances	Num	8	BEST.		Appearances
6	AvengersIntro	Num	8	MONYY7.		AvengersIntro
3	Current	Char	3	\$3.	\$3.	Current
8	Death1	Char	3	\$3.	\$3.	Death1
10	Death2	Char	3	\$3.	\$3.	Death2
12	Death3	Char	3	\$3.	\$3.	Death3
14	Death4	Char	3	\$3.	\$3.	Death4
16	Death5	Char	3	\$3.	\$3.	Death5
7	Honorary	Char	12	\$12.	\$12.	Honorary
1	Name	Char	35	\$35.	\$35.	Name
5	ProbationaryIntro	Num	8	MONYY7.		ProbationaryIntro
9	Return1	Char	3	\$3.	\$3.	Return1
11	Return2	Char	3	\$3.	\$3.	Return2
13	Return3	Char	3	\$3.	\$3.	Return3
15	Return4	Char	3	\$3.	\$3.	Return4
17	Return5	Char	3	\$3.	\$3.	Return5
4	Sex	Num	8	BEST.		Sex

2a)

Obs	Name	Appearances	Sex	AvengersIntro	YearIntro	Days	Years	Gender
1	Henry Jonathan "Hank" Pym	22000	0	SEP1963	1963	21215	58.08	Male
2	Janet van Dyne	1165	1	SEP1963	1963	21215	58.08	Female
3	Anthony Edward "Tony" Stark	3068	0	SEP1963	1963	21215	58.08	Male
4	Robert Bruce Banner	2089	0	SEP1963	1963	21215	58.08	Male
5	Thor Odinson	2402	0	SEP1963	1963	21215	58.08	Male
6	Richard Milhouse Jones	612	0	SEP1963	1963	21215	58.08	Male
7	Steven Rogers	3458	0	MAR1964	1964	21033	57.59	Male
8	Clinton Francis Barton	1456	0	MAY1965	1965	20607	56.42	Male
9	Pietro Maximoff	769	0	MAY1965	1965	20607	56.42	Male
10	Wanda Maximoff	1214	1	MAY1965	1965	20607	56.42	Female
11	Jacques Duquesne	115	0	SEP1965	1965	20484	56.08	Male
12	Heracles	741	0	OCT1967	1967	19724	54.00	Male
13	T'Challa	780	0	MAY1968	1968	19511	53.42	Male
14	Victor Shade (alias)	1036	0	NOV1968	1968	19327	52.91	Male
15	Dane Whitman	482	0	DEC1969	1969	18932	51.83	Male
16	Natalia Alianovna Romanova	1112	1	MAY1973	1973	17685	48.42	Female
17	Brandt	160	1	AUG1973	1973	17593	48.17	Female
18	Henry P. McCoy	1886	0	SEP1976	1976	16466	45.08	Male
19	Heather Douglas	332	1	SEP1976	1976	16466	45.08	Female
20	Patsy Walker	557	1	SEP1976	1976	16466	45.08	Female
21	Matthew Liebowitz (birth name)	197	0	AUG1975	1975	16863	46.17	Male
22	Robert L. Frank Sr.	106	0	JUL1978	1978	15798	43.25	Male
23	Simon Williams	692	0	MAR1979	1979	15555	42.59	Male
24	Yondu Udonta	109	0	FEB1978	1978	15948	43.66	Male
25	Martinex T'Naga	100	0	FEB1978	1978	15948	43.66	Male
26	Charlie-27	132	0	FEB1978	1978	15948	43.66	Male
27	Nichollette Gold	108	1	FEB1978	1978	15948	43.66	Female
28	Stakar	100	0	FEB1978	1978	15948	43.66	Male
29	Vance Astrovik	156	0	FEB1978	1978	15948	43.66	Male
30	Mar-Vell	254	0	JUL1978	1978	15798	43.25	Male
31	Carol Susan Jane Danvers	935	1	APR1979	1979	15524	42.50	Female

2b)

Avenger Character History

Obs	Character Name	Gender	Number of Appearance	Date Introduced	Year Introduced	Years Since Introduction
1	Henry Jonathan "Hank" Pym	Male	22,000	09/01/63	1963	58.08
2	Janet van Dyne	Female	1,165	09/01/63	1963	58.08
3	Anthony Edward "Tony" Stark	Male	3,068	09/01/63	1963	58.08
4	Robert Bruce Banner	Male	2,089	09/01/63	1963	58.08
5	Thor Odinson	Male	2,402	09/01/63	1963	58.08
6	Richard Milhouse Jones	Male	612	09/01/63	1963	58.08
7	Steven Rogers	Male	3,458	03/01/64	1964	57.59
8	Clinton Francis Barton	Male	1,456	05/01/65	1965	56.42
9	Pietro Maximoff	Male	769	05/01/65	1965	56.42
10	Wanda Maximoff	Female	1,214	05/01/65	1965	56.42
11	Jacques Duquesne	Male	115	09/01/65	1965	56.08
12	Heracles	Male	741	10/01/67	1967	54.00
13	T'Challa	Male	780	05/01/68	1968	53.42
14	Victor Shade (alias)	Male	1,036	11/01/68	1968	52.91
15	Dane Whitman	Male	482	12/01/69	1969	51.83
16	Natalia Alianovna Romanova	Female	1,112	05/01/73	1973	48.42
17	Brandt	Female	160	08/01/73	1973	48.17
18	Henry P. McCoy	Male	1,886	09/01/76	1976	45.08
19	Heather Douglas	Female	332	09/01/76	1976	45.08
20	Patsy Walker	Female	557	09/01/76	1976	45.08
21	Matthew Liebowitz (birth name)	Male	197	08/01/75	1975	46.17
22	Robert L. Frank Sr.	Male	106	07/01/78	1978	43.25
23	Simon Williams	Male	692	03/01/79	1979	42.59
24	Yondu Udonta	Male	109	02/01/78	1978	43.66
25	Martinex T'Naga	Male	100	02/01/78	1978	43.66
26	Charlie-27	Male	132	02/01/78	1978	43.66
27	Nichollette Gold	Female	108	02/01/78	1978	43.66
28	Stakar	Male	100	02/01/78	1978	43.66

2d) Add Format Gender \$6. j

2f)

Who's the Favorite Avenger

Popularity	Character Name	Date Introduced
High	Henry Jonathan "Hank" Pym	01SEP1963
	Janet van Dyne	01SEP1963
	Anthony Edward "Tony" Stark	01SEP1963
	Robert Bruce Banner	01SEP1963
	Thor Odinson	01SEP1963
	Steven Rogers	01MAR1964
	Clinton Francis Barton	01MAY1965
	Wanda Maximoff	01MAY1965
	Victor Shade (alias)	01NOV1968
	Natalia Alianovna Romanova	01MAY1973
	Henry P. McCoy	01SEP1976
	Benjamin Jacob Grimm	01JUN1986
	Namor McKenzie	01DEC1985
	Reed Richards	01FEB1989
	Susan Richards (nee Storm)	01FEB1989
	Peter Benjamin Parker	01APR1990
	James "Logan" Howlett	01JUN2005
	Doctor Stephen Vincent Strange	01FEB2007
	Ororo Munroe	01JAN2012
	Matt Murdock	01NOV2011
N = 20		

Who's the Favorite Avenger

Popularity	Character Name	Date Introduced
low	Moirá Brandon	01NOV1993
	Bonita Juarez	01SEP1987
	Marrina Smallwood	01DEC1987
	Ravonna Lexus Renslayer	01MAY1988
	Rita DeMara	01NOV1988
	Dennis Dunphy	01JAN1988
	Ashley Crawford	01JUL1989
	DeMarr Davis	01JUL1989
	Val Ventura	01JUL1989
	Craig Hollis	01JUL1989
	Gene Lorrene	01APR2005
	Dorreen Green	01MAY2005
	Doug Taggart	01MAY2005
	Greg Willis	01AUG2009
	Phillip Javert	01DEC1992
	Anthony Edward Stark	01FEB1996
	"Giulietta Nefaria"	01APR1996
	María de Guadalupe Santiago	01JUL2000
	Kelsey Leigh Shorr	01JUN2004
	Maya Lopez	01NOV2005
	Nathaniel Richards	01APR2005
	Thomas "Tommy" Shepherd	01JUN2006
	Eric O'Grady	01MAY2010
	Dennis Sykes	01SEP2010
	John Aman	01DEC2010