Exercise 1a

The CONTENTS Procedure

Data Set Name	HW02.FMLI143	**Observations**	4879
Member Type	DATA	**Variables**	30
Engine	V9	**Indexes**	0
Created	02/13/2019 17:21:58	**Observation Length**	144
Last Modified	02/13/2019 17:21:58	**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)		

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format
3	AGE2	Num	8	2.
2	AGE_REF	Num	8	2.
4	BATHRMQ	Num	8	3.
5	BEDROOMQ	Num	8	3.
6	BLS_URBN	Char	1	$1.
29	CUID	Num	8	7.
8	EDUCA2	Char	2	$2.
7	EDUC_REF	Char	2	$2.
9	FAM_SIZE	Num	8	2.
10	FINCATAX	Num	8	9.
11	FINCBTAX	Num	8	9.
25	HHID	Num	8	3.
24	HH_CU_Q	Num	8	2.
12	HLFBATHQ	Num	8	3.
26	INCLASS	Char	2	$2.
28	INCLASS2	Char	1	$1.
30	INTERI	Num	8	1.
13	MARITAL1	Char	1	$1.
1	NEWID	Num	8	8.
14	NO_EARNR	Num	8	2.
15	NUM_AUTO	Num	8	2.

Exercise 1a

The CONTENTS Procedure

Alphabetic List of Variables and Attributes				
#	**Variable**	**Type**	**Len**	**Format**
16	PRINEARN	Char	2	$2.
17	QINTRVMO	Char	2	$2.
18	QINTRVYR	Char	4	$4.
19	RACE2	Char	1	$1.
20	REF_RACE	Char	1	$1.
21	REGION	Char	1	$1.
23	SEX2	Char	1	$1.
22	SEX_REF	Char	1	$1.
27	STATE	Char	2	$2.

Exercise 1a

The CONTENTS Procedure

Data Set Name	HW02.MEMI143	**Observations**	12032
Member Type	DATA	**Variables**	9
Engine	V9	**Indexes**	0
Created	02/13/2019 17:21:58	**Observation Length**	40
Last Modified	02/13/2019 17:21:58	**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)		

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format
2	AGE	Num	8	2.
3	CU_CODE	Char	1	$1.
4	EDUCA	Char	1	$1.
5	MARITAL	Char	1	$1.
6	MEMBNO	Num	8	2.
9	MEMBRACE	Char	1	$1.
1	NEWID	Num	8	8.
7	SALARYX	Num	8	10.
8	SEX	Char	1	$1.

Exercise 1b

The CONTENTS Procedure

Data Set Name	HW02.FMLI143_JOSHLJ2	**Observations**	4879
Member Type	DATA	**Variables**	30
Engine	V9	**Indexes**	0
Created	09/30/2019 16:47:57	**Observation Length**	144
Last Modified	09/30/2019 16:47:57	**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)		

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Label
3	AGE2	Num	8	2.	Age of Spouse
2	AGE_REF	Num	8	2.	Age of Reference Person
4	BATHRMQ	Num	8	3.	# of Complete Bathrooms
5	BEDROOMQ	Num	8	3.	# Bedrooms
6	BLS_URBN	Char	1	$BLS_URBN.	Urban/Rural
29	CUID	Num	8	7.	Consumer unit identifying variable
8	EDUCA2	Char	2	$EDUC_REF.	Education of Spouse
7	EDUC_REF	Char	2	$EDUC_REF.	Education of Reference Person
9	FAM_SIZE	Num	8	2.	
10	FINCATAX	Num	8	9.	Amount of CU income after taxes
11	FINCBTAX	Num	8	9.	Amount of CU income before taxes
25	HHID	Num	8	3.	Identifier of household with more than one CU
24	HH_CU_Q	Num	8	2.	Count of CUs in household
12	HLFBATHQ	Num	8	3.	# Half Bathrooms
26	INCLASS	Char	2	$INCLASS.	Income class
28	INCLASS2	Char	1	$INCLASSTWO.	Income class based on INC_RANK
30	INTERI	Num	8	1.	Interview number
13	MARITAL1	Char	1	$MARITAL.	Marital Status
1	NEWID	Num	8	8.	CU identification number
14	NO_EARNR	Num	8	2.	Number of earners
15	NUM_AUTO	Num	8	2.	Number of owned automobiles

Exercise 1b

The CONTENTS Procedure

Alphabetic List of Variables and Attributes					
#	**Variable**	**Type**	**Len**	**Format**	**Label**
16	PRINEARN	Char	2	$2.	Member number of principal earner
17	QINTRVMO	Char	2	$QINTRVMO.	Interview month
18	QINTRVYR	Char	4	$4.	Interview year
19	RACE2	Char	1	$RACE.	Race of spouse
20	REF_RACE	Char	1	$RACE.	Race of reference person
21	REGION	Char	1	$REGION.	
23	SEX2	Char	1	$SEX.	Sex of spouse
22	SEX_REF	Char	1	$SEX.	Sex of reference person
27	STATE	Char	2	$STATE.	

Exercise 1b

The CONTENTS Procedure

Data Set Name	HW02.MEMI143_JOSHLJ2	**Observations**	12032
Member Type	DATA	**Variables**	9
Engine	V9	**Indexes**	0
Created	09/30/2019 16:47:57	**Observation Length**	40
Last Modified	09/30/2019 16:47:57	**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)		

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Label
2	AGE	Num	8	2.	
3	CU_CODE	Char	1	$CU_CODE.	Member relationship to reference person
4	EDUCA	Char	1	$EDUCA.	Education of member
5	MARITAL	Char	1	$MARITAL.	Marital Status
6	MEMBNO	Num	8	2.	Person line number
9	MEMBRACE	Char	1	$RACE.	Race of member
1	NEWID	Num	8	8.	CU identification number
7	SALARYX	Num	8	10.	Amount of wages or salary income received
8	SEX	Char	1	$SEX.	

Exercise 1c

Obs	NEWID	CUID	AGE_REF	BLS_URBN	MARITAL1	FINCATAX
1	2750024	275002	70	Urban	Widowed	0
2	2750074	275007	63	Urban	Married	1403749
3	2750094	275009	75	Urban	Widowed	11119
4	2750124	275012	47	Urban	Married	65300
5	2750134	275013	59	Urban	Separated	46930
6	2750164	275016	31	Urban	Married	132950
7	2750174	275017	60	Rural	Married	0
8	2750194	275019	61	Urban	Widowed	7993
9	2750214	275021	60	Urban	Divorced	0
10	2750244	275024	78	Urban	Separated	41838

Exercise 1c

Obs	NEWID	CU_CODE	MARITAL	SALARYX
1	2874002	Reference person	Never Married	102000
2	2874012	Reference person	Married	44832
3	2874012	Spouse	Married	17500
4	2874012	Child or adopted child	Never Married	.
5	2874012	Child or adopted child	Never Married	.
6	2874042	Reference person	Never Married	.
7	2874042	Child or adopted child	Never Married	.
8	2874042	Child or adopted child	Never Married	.
9	2874062	Reference person	Widowed	117000
10	2874112	Reference person	Married	.

Exercise 1f

The FREQ Procedure

Table of BEDROOMQ by FAM_SIZE														
BEDROOMQ(# Bedrooms)	**FAM_SIZE**													
Frequency Percent Row Pct Col Pct	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**13**	**14**	**Total**
0	30 0.62 83.33 2.05	4 0.08 11.11 0.26	2 0.04 5.56 0.29	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	36 0.75
1	394 8.17 66.67 26.91	133 2.76 22.50 8.58	36 0.75 6.09 5.14	14 0.29 2.37 2.37	11 0.23 1.86 3.51	2 0.04 0.34 1.56	0 0.00 0.00 0.00	1 0.02 0.17 5.56	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	591 12.26
2	456 9.46 37.13 31.15	433 8.98 35.26 27.94	182 3.77 14.82 25.96	107 2.22 8.71 18.14	33 0.68 2.69 10.54	13 0.27 1.06 10.16	3 0.06 0.24 6.82	0 0.00 0.00 0.00	1 0.02 0.08 12.50	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1228 25.47
3	443 9.19 23.21 30.26	673 13.96 35.25 43.42	317 6.57 16.61 45.22	272 5.64 14.25 46.10	134 2.78 7.02 42.81	45 0.93 2.36 35.16	16 0.33 0.84 36.36	6 0.12 0.31 33.33	2 0.04 0.10 25.00	0 0.00 0.00 0.00	1 0.02 0.05 50.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1909 39.59
4	112 2.32 13.24 7.65	256 5.31 30.26 16.52	144 2.99 17.02 20.54	152 3.15 17.97 25.76	107 2.22 12.65 34.19	47 0.97 5.56 36.72	15 0.31 1.77 34.09	8 0.17 0.95 44.44	3 0.06 0.35 37.50	1 0.02 0.12 50.00	1 0.02 0.12 50.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	846 17.54
5	21 0.44 12.14 1.43	43 0.89 24.86 2.77	18 0.37 10.40 2.57	38 0.79 21.97 6.44	23 0.48 13.29 7.35	17 0.35 9.83 13.28	7 0.15 4.05 15.91	2 0.04 1.16 11.11	2 0.04 1.16 25.00	1 0.02 0.58 50.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.02 0.58 100.00	173 3.59
6	7 0.15 21.88 0.48	7 0.15 21.88 0.45	1 0.02 3.13 0.14	5 0.10 15.63 0.85	5 0.10 15.63 1.60	3 0.06 9.38 2.34	2 0.04 6.25 4.55	1 0.02 3.13 5.56	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.02 3.13 100.00	0 0.00 0.00 0.00	32 0.66
7	1 0.02 14.29 0.07	1 0.02 14.29 0.06	1 0.02 14.29 0.14	2 0.04 28.57 0.34	0 0.00 0.00 0.00	1 0.02 14.29 0.78	1 0.02 14.29 2.27	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	7 0.15
Total	1464 30.36	1550 32.14	701 14.54	590 12.24	313 6.49	128 2.65	44 0.91	18 0.37	8 0.17	2 0.04	2 0.04	1 0.02	1 0.02	4822 100.00
Frequency Missing = 57														

Looking at the frequency table above, we can see there is some correlation between the family size and number of bedrooms. You can see that as the family size increases, so does the number of bedrooms. If you look at number of bedrooms, we can see that as # of bedrooms increase there are more counts for larger family sizes. Overall, there is correlation between these variables that can be seen.

Exercise 1f

The FREQ Procedure

Table of FAM_SIZE by BATHRMQ									
FAM_SIZE	**BATHRMQ(# of Complete Bathrooms)**								
Frequency Percent Row Pct Col Pct	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
1	10 0.21 0.68 66.67	921 19.10 62.91 41.69	456 9.46 31.15 21.83	65 1.35 4.44 15.55	12 0.25 0.82 16.22	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1464 30.36
2	1 0.02 0.06 6.67	646 13.40 41.65 29.24	739 15.33 47.65 35.38	134 2.78 8.64 32.06	23 0.48 1.48 31.08	5 0.10 0.32 41.67	1 0.02 0.06 50.00	2 0.04 0.13 66.67	1551 32.17
3	2 0.04 0.29 13.33	288 5.97 41.08 13.04	333 6.91 47.50 15.94	67 1.39 9.56 16.03	9 0.19 1.28 12.16	2 0.04 0.29 16.67	0 0.00 0.00 0.00	0 0.00 0.00 0.00	701 14.54
4	1 0.02 0.17 6.67	207 4.29 35.08 9.37	275 5.70 46.61 13.16	86 1.78 14.58 20.57	16 0.33 2.71 21.62	4 0.08 0.68 33.33	0 0.00 0.00 0.00	1 0.02 0.17 33.33	590 12.24
5	1 0.02 0.32 6.67	89 1.85 28.43 4.03	177 3.67 56.55 8.47	36 0.75 11.50 8.61	8 0.17 2.56 10.81	1 0.02 0.32 8.33	1 0.02 0.32 50.00	0 0.00 0.00 0.00	313 6.49
6	0 0.00 0.00 0.00	38 0.79 29.92 1.72	67 1.39 52.76 3.21	19 0.39 14.96 4.55	3 0.06 2.36 4.05	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	127 2.63
7	0 0.00 0.00 0.00	12 0.25 27.27 0.54	23 0.48 52.27 1.10	6 0.12 13.64 1.44	3 0.06 6.82 4.05	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	44 0.91
8	0 0.00 0.00 0.00	4 0.08 22.22 0.18	12 0.25 66.67 0.57	2 0.04 11.11 0.48	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	18 0.37
9	0 0.00 0.00 0.00	3 0.06 37.50 0.14	3 0.06 37.50 0.14	2 0.04 25.00 0.48	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	8 0.17
10	0 0.00 0.00 0.00	0 0.00 0.00 0.00	2 0.04 100.00 0.10	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	2 0.04

Exercise 1f

The FREQ Procedure

Table of FAM_SIZE by BATHRMQ									
FAM_SIZE	**BATHRMQ(# of Complete Bathrooms)**								
Frequency Percent Row Pct Col Pct	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
11	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.02 50.00 0.05	1 0.02 50.00 0.24	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	2 0.04
13	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.02 100.00 0.05	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.02
14	0 0.00 0.00 0.00	1 0.02 100.00 0.05	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.02
Total	15 0.31	2209 45.81	2089 43.32	418 8.67	74 1.53	12 0.25	2 0.04	3 0.06	4822 100.00
Frequency Missing = 57									

Looking at the frequency table above of # of complete bathrooms and family size, we can see that there is some correlation between them as the number of bathrooms tends to be higher for larger family sizes. However, it seems that regardless of family size most families have between 1-2 bathrooms as that is the most common. Overall, there is some correlation between these variables although not much and there seems to be a common number of bathrooms regardless of family size.

Exercise 1g

The MEANS Procedure

Analysis Variable : SALARYX Amount of wages or salary income received						
Education of member	**N Obs**	**Mean**	**Median**	**Quartile Range**	**Lower 95% CL for Mean**	**Upper 95% CL for Mean**
No schooling completed, or less than 1 year	44	17889.60	19000.00	3872.00	6449.00	29330.20
Nursery, kindergarten, and elementary (grades 1-8)	477	21394.28	20000.00	20200.00	18432.00	24356.56
High school (grades 9-12, no degree)	1416	19490.19	16000.00	26500.00	17347.88	21632.50
High school graduate – high school diploma or the equivalent (GED)	2496	31688.00	25000.00	27800.00	29789.84	33586.16
Some college but no degree	1949	34278.82	27000.00	35300.00	31839.69	36717.95
Associate's degree in college	740	42909.50	37000.00	36500.00	39153.83	46665.17
Bachelor's degree (BA, AB, BS, etc.)	1710	58523.84	47000.00	48650.00	55009.62	62038.06
Master's professional, or doctorate degree (MA, MS, MBA, MD, JD, PhD, etc.)	977	87125.75	69000.00	67000.00	80534.55	93716.94

As we can see above in the means procedure between education and the salary, there is a connection between the two. We can see that as education increases in an individual, the mean, median, and quartile range increase in salary. We can also see that our 95% confidence intervals for mean salaries increase with education. Overall, as education increases so does salary.

Exercise 1h

The MEANS Procedure

Analysis Variable : SALARYX Amount of wages or salary income received						
AGE	**N Obs**	**Mean**	**Median**	**Quartile Range**	**Lower 95% CL for Mean**	**Upper 95% CL for Mean**
Child	1633	.	.	.	.	.
Teens	1569	6147.29	2500.00	7150.00	4562.11	7732.48
Twenties	1605	26128.32	20000.00	27400.00	24495.14	27761.49
Thirties	1545	48508.40	38000.00	41000.00	45497.27	51519.53
Forties	1474	57725.85	42900.00	49000.00	53575.14	61876.56
Fifties	1740	56704.37	42000.00	46000.00	52929.28	60479.45
Sixties	1326	45116.73	31784.00	45000.00	40561.23	49672.23
Seventies	738	35251.94	13000.00	37200.00	21857.54	48646.35
Eighties	402	39293.20	10500.00	28464.00	-17613.84	96200.24

Looking at the means procedure above between age and salary, where age is formatted to display the age range of the individual, we can see that there is a definite connection between the variables. Starting from children, mean salary continues to rise until the individual reaches their sixties, in which their mean salary falls (excluding the increase from seventies to eighties). This falls in line with how salaries should increase with age, as children do not make money, teens make some money, then as you get into your twenties your salary increases until you retire. Once retired, your salary decreases. In our data, we can see that those in their eighties have a higher salary than those in their seventies. I believe this is due to a smaller sample/those in their eighties just having a larger retirement fund.

Looking at the quartile ranges, we can see that this follows a similar output of the mean. For confidence intervals, we can see that as age increases until fifties our salary estimate increases. After fifties, our estimates decrease. Overall, there is a clear connection between age and salary.