

QUESTION ONE.

Part A. Cross-tab of mother's work status and the number of kids she has: cell counts.

<u>Number of Kids</u>	Mother is:	
	<u>Not Working</u>	<u>Working</u>
1	119	73
2 or 3	2,150	2,218
4 or More	476	473

Part B. Cross-tab of mother's work status and the age of her youngest child: cell counts and row percentages.

<u>Age of Youngest Child</u>	Mother is:	
	<u>Not Working</u>	<u>Working</u>
Under 3 years old	1,489 58.12%	1,073 41.88%
3 to 5 years old	1,026 45.22%	1,243 54.78%
Over 6 years old	230 33.92%	448 66.08%

Part C. Percent of working mothers by standard metropolitan area.

<u>SMA</u>	<u>Employment Rate</u>
Boston	52.6%
Brockton	60.7%
Lowell	55.1%
New Bedford	45.1%
Providence	57.5%
Springfield	42.9%

Part D. Table of the centiles of the percent of group day care providers that are NAEYC accredited.

<u>Centile</u>	<u>Percent Acc.</u>
10	0.000
20	0.000
30	0.051
40	0.128
50	0.164
60	0.208
70	0.230
80	0.253
90	0.412

Part E. Average wage and average weekly hours by education level.

	No High School Deg.	High School Degree	Some College	College
Avg. Weekly Wage	\$7.13	\$7.49	\$7.86	\$8.18
Avg. Weekly Hrs. Worked	31.98	32.53	31.92	33.21

QUESTION TWO.

Part A.

. describe;

Contains data from cps\_may\_2006\_workers.dta

obs: 3,175  
vars: 20 5 Sep 2008 07:43  
size: 266,700 (97.5% of memory free)

---

storage	display	value		
variable	name	type	format	label
wage		float	%9.0g	Hourly wage
lnwage		float	%9.0g	Ln(wage)
status		float	%9.0g	marstats Marital Status
married		float	%9.0g	Married=1; 0 otherwise
single		float	%9.0g	Single=1; 0 otherwise
divorced		float	%9.0g	Divorced=1; 0 otherwise
sex		float	%9.0g	sexes Sex: 0=Female, 1=Male
male		float	%9.0g	Male=1; Female=0
female		float	%9.0g	Female=1; Male=0
educ		float	%9.0g	edclass Education Class
nohsd		float	%9.0g	Educ Class: No HS Degree=1
hsd		float	%9.0g	Educ Class: High Schl Deg=1
somocol		float	%9.0g	Educ Class: Some College=1
college		float	%9.0g	Educ Class: College=1
race		float	%9.0g	races Race Class
white		float	%9.0g	Race: White=1; 0 otherwise
black		float	%9.0g	Race: Black=1; 0 otherwise
natam		float	%9.0g	Race: Native American=1; 0 o/w
asian		float	%9.0g	Race: Asian=1; 0 otherwise
othrc		float	%9.0g	Race: Other Race=1; 0 o/w

---

. sum;

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	3175	14.33704	7.746619	2.1	75
lnwage	3175	2.549565	.464183	.7419373	4.317488
status	3175	1.737323	.8439949	1	3
married	3175	.5218898	.4995993	0	1
single	3175	.2592126	.4382715	0	1
divorced	3175	.2188976	.4135642	0	1
sex	3175	.48	.4996785	0	1
male	3175	.48	.4996785	0	1

female	3175	.52	.4996785	0	1
educ	3175	2.507717	.9115863	1	4
nohsd	3175	.1376378	.3445737	0	1
hsd	3175	.3697638	.4828167	0	1
somecol	3175	.3398425	.4737302	0	1
college	3175	.1527559	.3598087	0	1
race	3175	1.326614	.8329806	1	5
white	3175	.8107087	.3918016	0	1
black	3175	.1275591	.3336507	0	1
natam	3175	.0100787	.0999015	0	1
asian	3175	.0277165	.1641853	0	1
othrc	3175	.023937	.152877	0	1

\* Part B.

The fact that the average value of Sex is 0.48 means that 51% of the sample are males and 48% are females. To convince yourself, you should enter

```
tab sex
```

followed by

```
tab sex, nol
```

to see that males are associated with the value of 0 and females are associated with the value of 1. Once it is known that the variable equals 1 for females and equals 0 for males, then the average is the percent of the sample with a value of 1.

\* Part C.

```
. sum wage if wage>=5 & wage<=15;
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	2132	10.49675	2.603536	5	15

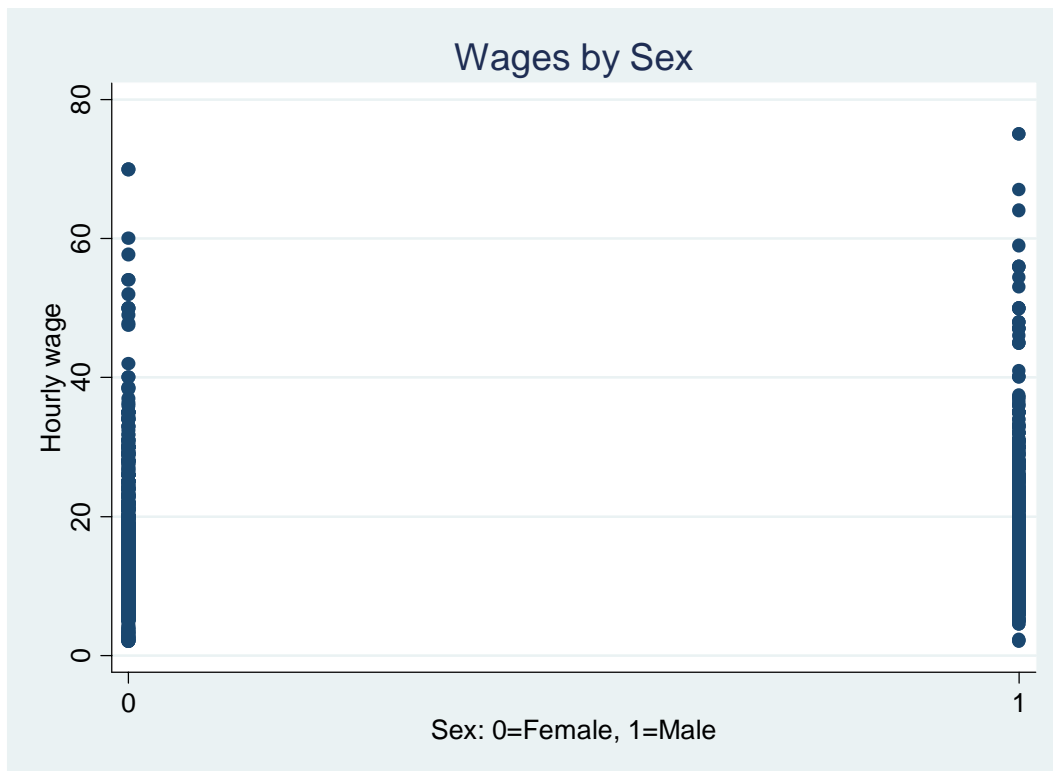
Thus, 2,132 of the 3,175 observations are from individuals with a wage of at least \$5 per hour and at most \$15 per hour.

\* Part D.

```
bysort educ: sum wage;
```

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
-> educ = No HSD					
wage	437	10.72929	4.214949	2.13	32
-> educ = HSD					
wage	1174	13.41049	6.084431	2.1	60
-> educ = Some Col					
wage	1079	14.46709	6.988892	2.1	56
-> educ = Col Grad					
wage	485	19.54122	11.74781	2.13	75
-----+-----					

Part E.



This graph is fairly meaningless, because all of the wage numbers are plotted on top of one another for men and women.

Part F.



According to the graph, it looks like the average female wage is about \$13 per hour, while the average male wage is about \$15.25 per hour.

Part G.



Women appear to have a lower distribution of wages than men at all points. This is seen by women having a greater density at lower wages.

PROGRAM FOR STATA PROJECT ONE

```
#delimit;  
set more 1;
```

```
* Rob Lemke;  
* September 14, 2008;
```

```
log using project1.log, replace;
```

```
* PROBLEM ONE;
```

```
use welfare_edited.dta, replace;
```

```
* part A;  
gen kids=2-(noofdeps==1)+(noofdeps>=4);  
label define kids 1 "1 Kid" 2 "2or3Kids" 3 "4+ Kids";  
label values kids kids;  
label variable kids "Number of kids: 1, 2 or 3, 4+";  
tab kids working;
```

```
* part B;  
gen kidage=2-(ageych<=2)+(ageych>=6);  
label define kidages 1 "Under 3" 2 "3 to 5" 3 "Over 5";  
label values kidage kidages;  
label variable kidage "Age of yngst kid: 0-2,3-5,6+";  
tab kidage working, row;
```

```
* part C;  
sort smsa;  
by smsa: sum working;
```

```
* part D;  
_pctile gdcnaeyc, p(10 20 30 40 50 60 70 80 90);  
gen p10=r(r1);  
gen p20=r(r2);  
gen p30=r(r3);  
gen p40=r(r4);  
gen p50=r(r5);  
gen p60=r(r6);  
gen p70=r(r7);  
gen p80=r(r8);  
gen p90=r(r9);  
sum p10 p20 p30 p40 p50 p60 p70 p80 p90;
```

```
* part E;  
sort educ;  
by educ: sum wage hours if working==1;  
clear;
```

```
* THE COLLAPSE COMMAND;  
* Consider the following commands that could have also been  
  used to complete parts C, D, and E.;
```

```
* part C redux;  
use welfare_edited.dta, replace;
```

```

sort smsa;
collapse (mean) working, by(smsa);
list smsa working, clean;
clear;

* part D redux;
use welfare_edited.dta, replace;
sort gdcnaeyc;
collapse (p10) p10=gdcnaeyc (p20) p20=gdcnaeyc (p30) p30=gdcnaeyc
          (p40) p40=gdcnaeyc (p50) p50=gdcnaeyc (p60) p60=gdcnaeyc
          (p70) p70=gdcnaeyc (p80) p80=gdcnaeyc (p90) p90=gdcnaeyc;
list, clean;
clear;

* part D redux redux;
use welfare_edited.dta, replace;
centile gdcnaeyc, centile(10 20 30 40 50 60 70 80 90);
clear;

* part E;
use welfare_edited.dta, replace;
sort educ;
collapse (mean) wage hours, by(educ);
list educ wage hours, clean;
clear;

* PROBLEM TWO;

use cps_may_2006_workers.dta, replace;

* part A;
label variable wage      "Hourly wage";
label variable lnwage    "Ln(wage)";
label variable status    "Marital Status";
label variable married   "Married=1; 0 otherwise";
label variable single    "Single=1; 0 otherwise";
label variable divorced  "Divorced=1; 0 otherwise";
label variable sex       "Sex: 0=Female, 1=Male";
label variable male      "Male=1; Female=0";
label variable female    "Female=1; Male=0";
label variable educ      "Education Class";
label variable nohsd     "Educ Class: No HS Degree=1";
label variable hsd       "Educ Class: High Schl Deg=1";
label variable somecol   "Educ Class: Some College=1";
label variable college   "Educ Class: College=1";
label variable race      "Race Class";
label variable white     "Race: White=1; 0 otherwise";
label variable black     "Race: Black=1; 0 otherwise";
label variable natam     "Race: Native American=1; 0 o/w";
label variable asian     "Race: Asian=1; 0 otherwise";
label variable othrc     "Race: Other Race=1; 0 o/w";
describe;
sum;

* part B;
* writing;

```

```

* part C;
sum wage if wage>=5 & wage<=15;

* part D;
bysort educ: sum wage;

* part E;
scatter wage sex, title("Wages by Sex") xlabel(0 1) saving(partE.gph,
replace);

* part F;
graph bar (mean) wage, over(sex) title("Wages by Sex") saving(partF.gph,
replace);

* part G;
hist wage if male==1, bin(50) title(Men) saving(wages_men.gph, replace);
hist wage if male==0, bin(50) title(Women) saving(wages_women.gph, replace);
graph combine wages_men.gph wages_women.gph, ti("Wages by Sex")
saving(partG.gph, replace) ycommon;

clear;
log close;

```

LOG FILE FOR STATA PROJECT ONE

```

log: C:\a_office\Econ 330\Stata\projects\project1\project1.log
log type: text
opened on: 14 Sep 2008, 16:29:55

```

```

. * PROBLEM ONE;
. use welfare_edited.dta, replace;

. * part A;
. gen kids=2-(noofdeps==1)+(noofdeps>=4);

. label define kids 1 "1 Kid" 2 "2or3Kids" 3 "4+ Kids";

. label values kids kids;

. label variable kids "Number of kids: 1, 2 or 3, 4+";

. tab kids working;

```

kids: 1, 2 or 3, 4+	=1 if working; =0 if not working		Total
	0	1	
1 Kid	119	73	192
2or3Kids	2,150	2,218	4,368
4+ Kids	476	473	949
Total	2,745	2,764	5,509

```

. * part B;
. gen kidage=2-(ageych<=2)+(ageych>=6);

. label define kidages 1 "Under 3" 2 "3 to 5" 3 "Over 5";

. label values kidage kidages;

. label variable kidage "Age of yngst kid: 0-2,3-5,6+";

. tab kidage working, row;

```

```

+-----+
| Key   |
+-----+
|       |
| frequency |
| row percentage |
+-----+

```

Age of yngst kid: 0-2,3-5,6+	=1 if working; =0 if not working		Total
	0	1	
Under 3	1,489 58.12	1,073 41.88	2,562 100.00
3 to 5	1,026 45.22	1,243 54.78	2,269 100.00
Over 5	230 33.92	448 66.08	678 100.00
Total	2,745 49.83	2,764 50.17	5,509 100.00

```

. * part C;
. sort smsa;

. by smsa: sum working;

```

-> smsa = Boston

Variable	Obs	Mean	Std. Dev.	Min	Max
working	2808	.5259972	.4994126	0	1

-> smsa = Brockton

Variable	Obs	Mean	Std. Dev.	Min	Max
working	28	.6071429	.4973475	0	1

-> smsa = Lowell

Variable	Obs	Mean	Std. Dev.	Min	Max
working	443	.5507901	.497976	0	1

-> smsa = New Bedford

Variable	Obs	Mean	Std. Dev.	Min	Max
working	501	.4510978	.4981002	0	1

-> smsa = Providence

Variable	Obs	Mean	Std. Dev.	Min	Max
working	402	.5746269	.4950155	0	1

-> smsa = Springfield

Variable	Obs	Mean	Std. Dev.	Min	Max
working	1327	.4287867	.4950893	0	1

```
. * part D;  
. _pctile gdcnaeyc, p(10 20 30 40 50 60 70 80 90);  
  
. gen p10=r(r1);  
. gen p20=r(r2);  
. gen p30=r(r3);  
. gen p40=r(r4);  
. gen p50=r(r5);  
. gen p60=r(r6);  
. gen p70=r(r7);  
. gen p80=r(r8);  
. gen p90=r(r9);  
  
. sum p10 p20 p30 p40 p50 p60 p70 p80 p90;
```

Variable	Obs	Mean	Std. Dev.	Min	Max
p10	5509	0	0	0	0
p20	5509	0	0	0	0
p30	5509	.051	0	.051	.051
p40	5509	.128	0	.128	.128
p50	5509	.164	0	.164	.164
p60	5509	.208	0	.208	.208
p70	5509	.23	0	.23	.23
p80	5509	.253	0	.253	.253
p90	5509	.412	0	.412	.412

```
. * part E;
. sort educ;

. by educ: sum wage hours if working==1;
```

```
-> educ = College Degree
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	106	8.177075	2.310155	3.5	17
hours	106	33.20755	8.53644	5	40

```
-> educ = HS Degree
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	1411	7.485145	2.311163	2.55	50
hours	1411	32.52729	8.430197	3	60

```
-> educ = No HS Degree
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	907	7.133638	1.829805	2	13.45
hours	907	31.98346	8.444485	5	52

```
-> educ = Some College
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	340	7.860147	2.117597	2.15	16
hours	340	31.90882	8.93275	5	55

```
. clear;
```

```
. * THE COLLAPSE COMMAND;
. * Consider the following commands that could have also been
> used to complete parts C, D, and E.;
```

```
. * part C redux;
. use welfare_edited.dta, replace;
```

```
. sort smsa;
```

```
. collapse (mean) working, by(smsa);
```

```
. list smsa working, clean;
```

	smsa	working
1.	Boston	.5259972
2.	Brockton	.6071429
3.	Lowell	.5507901
4.	New Bedford	.4510978
5.	Providence	.5746269
6.	Springfield	.4287867

```
. clear;
```

```

. * part D redux;
. use welfare_edited.dta, replace;

. sort gdcnaeyc;

. collapse (p10) p10=gdcnaeyc (p20) p20=gdcnaeyc (p30) p30=gdcnaeyc
>          (p40) p40=gdcnaeyc (p50) p50=gdcnaeyc (p60) p60=gdcnaeyc
>          (p70) p70=gdcnaeyc (p80) p80=gdcnaeyc (p90) p90=gdcnaeyc;

```

```

. list, clean;

```

```

      p10   p20   p30   p40   p50   p60   p70   p80   p90
1.      0     0   .051   .128   .164   .208   .23   .253   .412

```

```

. clear;

```

```

. * part D redux redux;
. use welfare_edited.dta, replace;

```

```

. centile gdcnaeyc, centile(10 20 30 40 50 60 70 80 90);

```

Variable	Obs	Percentile	Centile	-- Binom. Interp. -- [95% Conf. Interval]	
gdcnaeyc	5509	10	0	0	0
		20	0	0	0
		30	.051	.051	.115
		40	.128	.128	.144
		50	.164	.164	.166
		60	.208	.208	.208
		70	.23	.23	.23
		80	.253	.253	.253
		90	.412	.397	.412

```

. clear;

```

```

. * part E;
. use welfare_edited.dta, replace;

```

```

. sort educ;

```

```

. collapse (mean) wage hours, by(educ);

```

```

. list educ wage hours, clean;

```

```

      educ      wage      hours
1. College Degree  5.665164  23.00654
2.      HS Degree  4.226306  18.36575
3.   No HS Degree  2.92505   13.11438
4.   Some College  4.143333  16.82015

```

```

. clear;

```

```

. * PROBLEM TWO;
. use cps_may_2006_workers.dta, replace;

```

```

. * part A;
. label variable wage      "Hourly wage";

. label variable lnwage    "Ln(wage)";

. label variable status    "Marital Status";

. label variable married   "Married=1; 0 otherwise";

. label variable single    "Single=1; 0 otherwise";

. label variable divorced  "Divorced=1; 0 otherwise";

. label variable sex       "Sex: 0=Female, 1=Male";

. label variable male      "Male=1; Female=0";

.           label variable female    "Female=1; Male=0";

. label variable educ      "Education Class";

. label variable nohsd     "Educ Class: No HS Degree=1";

. label variable hsd       "Educ Class: High Schl Deg=1";

.           label variable somecol    "Educ Class: Some College=1";

.           label variable college    "Educ Class: College=1";

.           label variable race       "Race Class";

.           label variable white      "Race: White=1; 0 otherwise";

. label variable black     "Race: Black=1; 0 otherwise";

.           label variable natam      "Race: Native American=1; 0 o/w";

.           label variable asian      "Race: Asian=1; 0 otherwise";

.           label variable othrc      "Race: Other Race=1; 0 o/w";

. describe;

```

Contains data from cps\_may\_2006\_workers.dta

```

obs:      3,175
vars:      20                               5 Sep 2008 07:43
size:     266,700 (97.5% of memory free)

```

---

storage	display	value		
variable	name	type	format	label
wage		float	%9.0g	Hourly wage
lnwage		float	%9.0g	Ln(wage)
status		float	%9.0g	marstats Marital Status
married		float	%9.0g	Married=1; 0 otherwise
single		float	%9.0g	Single=1; 0 otherwise
divorced		float	%9.0g	Divorced=1; 0 otherwise

---

```

sex          float  %9.0g      sexes      Sex: 0=Female, 1=Male
male        float  %9.0g      Male=1; Female=0
female      float  %9.0g      Female=1; Male=0
educ        float  %9.0g      edclass    Education Class
nohsd       float  %9.0g      Educ Class: No HS Degree=1
hsd         float  %9.0g      Educ Class: High Schl Deg=1
somecol     float  %9.0g      Educ Class: Some College=1
college     float  %9.0g      Educ Class: College=1
race        float  %9.0g      races      Race Class
white       float  %9.0g      Race: White=1; 0 otherwise
black       float  %9.0g      Race: Black=1; 0 otherwise
natam       float  %9.0g      Race: Native American=1; 0 o/w
asian       float  %9.0g      Race: Asian=1; 0 otherwise
othrc       float  %9.0g      Race: Other Race=1; 0 o/w

```

Sorted by:

```
. sum;
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	3175	14.33704	7.746619	2.1	75
lnwage	3175	2.549565	.464183	.7419373	4.317488
status	3175	1.737323	.8439949	1	3
married	3175	.5218898	.4995993	0	1
single	3175	.2592126	.4382715	0	1
divorced	3175	.2188976	.4135642	0	1
sex	3175	.48	.4996785	0	1
male	3175	.48	.4996785	0	1
female	3175	.52	.4996785	0	1
educ	3175	2.507717	.9115863	1	4
nohsd	3175	.1376378	.3445737	0	1
hsd	3175	.3697638	.4828167	0	1
somecol	3175	.3398425	.4737302	0	1
college	3175	.1527559	.3598087	0	1
race	3175	1.326614	.8329806	1	5
white	3175	.8107087	.3918016	0	1
black	3175	.1275591	.3336507	0	1
natam	3175	.0100787	.0999015	0	1
asian	3175	.0277165	.1641853	0	1
othrc	3175	.023937	.152877	0	1

```

. * part B;
. * writing;
. * part C;
. sum wage if wage>=5 & wage<=15;

```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	2132	10.49675	2.603536	5	15

```
. * part D;
. bysort educ: sum wage;
```

```
-> educ = No HSD
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	437	10.72929	4.214949	2.13	32

```
-> educ = HSD
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	1174	13.41049	6.084431	2.1	60

```
-> educ = Some Col
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	1079	14.46709	6.988892	2.1	56

```
-> educ = Col Grad
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	485	19.54122	11.74781	2.13	75

```
. * part E;
. scatter wage sex, title("Wages by Sex") xlabel(0 1) saving(partE.gph,
replace);
(file partE.gph saved)
```

```
. * part F;
. graph bar (mean) wage, over(sex) title("Wages by Sex") saving(partF.gph,
replace);
(file partF.gph saved)
```

```
. * part G;
. hist wage if male==1, bin(50) title(Men) saving(wages_men.gph, replace);
(bin=50, start=2.130001, width=1.4574)
(file wages_men.gph saved)
```

```
. hist wage if male==0, bin(50) title(Women) saving(wages_women.gph,
replace);
(bin=50, start=2.0999999, width=1.358)
(file wages_women.gph saved)
```

```
. graph combine wages_men.gph wages_women.gph, ti("Wages by Sex")
saving(partG.gph, replace) ycommon;
(file partG.gph saved)
```

```
. clear;
. log close;
  log: C:\a_office\Econ 330\Stata\projects\project1\project1.log
  log type: text
closed on: 14 Sep 2008, 16:30:08
```