```
 1 . do "C:\Users\ahmad\AppData\Local\Temp\STD2968_000000.tmp"

 2 . clear

 3 . *use "C:\Users\ahmad\Desktop\Econ 435\term paper\2011Census\nhs2011_pumf.dta"
 4 .
 5 . *STATA is dropping values with scientific notations such as exponential to solve that we are go
 6 . import delimited "C:\Users\ahmad\Desktop\Econ 435\term paper\2011Census\nhs2011_pumf.tab", case
   (133 vars, 887,012 obs)

 7 . keep WAGES

 8 . rename WAGES WAGE

 9 . merge 1:1 _n using "C:\Users\ahmad\Desktop\Econ 435\term paper\2011Census\nhs2011_pumf.dta"

    Result                           # of obs.
    -----------------------------------------
    not matched                             0
    matched                           887,012  (_merge==3)
    -----------------------------------------

10 .
11 . *Dropping people who are self-employed
12 . drop if(COW>1)
   (423,029 observations deleted)

13 .
14 . *defining depended variable log function of wage
15 . drop if (WAGE== 8888888|WAGE== 9999999)
   (402 observations deleted)
```

```
16 .
17 . *Wage=1 not possible considering minimum wages, hence we will drop it,
18 . *it's probably because of age group of 15-17, which includes 15 year old(15 year old not allowe
19 . drop if(WAGE==1|WAGE==0)
(42,708 observations deleted)

20 .
21 . label variable WAGE "Wages of Individuals"

22 .
23 . *My sample is people who worked full time full year only.
24 . *I am controlling for number of weeks worked too, to avoid variations in wages due to number of
25 . *I cannot stress this enough, My sample includes people who Worked 49 to 52 weeks full time.
26 . *Alternatively, I could have used "drop if (FPTWK!=1)"
27 . *but this would have only accounted for full time and part time not number of hours/weeks
28 .
29 . drop if(WRKACT!=11)
(194,961 observations deleted)

30 .
31 . summ WAGE
```

Variable	Obs	Mean	Std. Dev.	Min	Max
WAGE	**225,912**	**60696.7**	**54771.89**	**1000**	**1074400**

```
32 .
33 . gen LWAGE=log(WAGE)

34 . label variable LWAGE "log function of wages of individuals"

35 . tab LWAGE
```

log function of wages of individuals	Freq.	Percent	Cum.
6.907755	**799**	**0.35**	**0.35**
7.600903	**468**	**0.21**	**0.56**
8.006368	**488**	**0.22**	**0.78**
8.294049	**472**	**0.21**	**0.99**
8.517193	**485**	**0.21**	**1.20**
8.699514	**551**	**0.24**	**1.44**
8.853665	**523**	**0.23**	**1.68**
8.987197	**576**	**0.25**	**1.93**
9.10498	**627**	**0.28**	**2.21**
9.21034	**866**	**0.38**	**2.59**
9.305651	**759**	**0.34**	**2.93**
9.392662	**980**	**0.43**	**3.36**
9.472705	**956**	**0.42**	**3.78**
9.546813	**1,045**	**0.46**	**4.25**
9.615806	**1,347**	**0.60**	**4.84**
9.680344	**1,339**	**0.59**	**5.44**
9.740969	**1,433**	**0.63**	**6.07**
9.798127	**1,803**	**0.80**	**6.87**
9.852194	**1,826**	**0.81**	**7.68**
9.903487	**2,380**	**1.05**	**8.73**
9.952278	**2,127**	**0.94**	**9.67**
9.998797	**2,267**	**1.00**	**10.68**
10.04325	**2,335**	**1.03**	**11.71**
10.08581	**2,430**	**1.08**	**12.78**
10.12663	**2,799**	**1.24**	**14.02**
10.16585	**2,598**	**1.15**	**15.17**
10.20359	**2,618**	**1.16**	**16.33**

10.23996	**2,726**	**1.21**	**17.54**
10.27505	**2,757**	**1.22**	**18.76**
10.30895	**3,580**	**1.58**	**20.34**
10.34174	**2,958**	**1.31**	**21.65**
10.37349	**3,247**	**1.44**	**23.09**
10.40426	**3,150**	**1.39**	**24.49**
10.43412	**3,254**	**1.44**	**25.93**
10.4631	**3,827**	**1.69**	**27.62**
10.49127	**3,608**	**1.60**	**29.22**
10.51867	**3,522**	**1.56**	**30.78**
10.54534	**3,607**	**1.60**	**32.37**
10.57132	**3,399**	**1.50**	**33.88**
10.59663	**4,335**	**1.92**	**35.80**
10.62133	**3,368**	**1.49**	**37.29**
10.64542	**3,670**	**1.62**	**38.91**
10.66896	**3,342**	**1.48**	**40.39**
10.69195	**3,138**	**1.39**	**41.78**
10.71442	**3,606**	**1.60**	**43.38**
10.7364	**3,122**	**1.38**	**44.76**
10.7579	**3,143**	**1.39**	**46.15**
10.77896	**3,384**	**1.50**	**47.65**
10.79958	**2,781**	**1.23**	**48.88**
10.81978	**3,866**	**1.71**	**50.59**
10.83958	**2,785**	**1.23**	**51.82**
10.859	**3,077**	**1.36**	**53.18**
10.87805	**2,757**	**1.22**	**54.40**
10.89674	**2,762**	**1.22**	**55.63**
10.91509	**2,993**	**1.32**	**56.95**
10.93311	**2,675**	**1.18**	**58.14**
10.95081	**2,606**	**1.15**	**59.29**
10.9682	**2,503**	**1.11**	**60.40**
10.98529	**2,347**	**1.04**	**61.44**
11.0021	**3,187**	**1.41**	**62.85**
11.01863	**2,206**	**0.98**	**63.82**
11.03489	**2,280**	**1.01**	**64.83**
11.05089	**2,265**	**1.00**	**65.84**
11.06664	**2,138**	**0.95**	**66.78**
11.08214	**2,459**	**1.09**	**67.87**
11.09741	**2,007**	**0.89**	**68.76**
11.11245	**1,948**	**0.86**	**69.62**
11.12726	**2,011**	**0.89**	**70.51**
11.14186	**1,885**	**0.83**	**71.35**
11.15625	**2,409**	**1.07**	**72.41**
11.17043	**1,914**	**0.85**	**73.26**
11.18442	**1,875**	**0.83**	**74.09**
11.19821	**1,840**	**0.81**	**74.90**
11.21182	**1,693**	**0.75**	**75.65**
11.22524	**1,935**	**0.86**	**76.51**
11.23849	**1,603**	**0.71**	**77.22**
11.25156	**1,584**	**0.70**	**77.92**
11.26446	**1,523**	**0.67**	**78.59**
11.2772	**1,422**	**0.63**	**79.22**
11.28978	**1,898**	**0.84**	**80.06**
11.3022	**1,393**	**0.62**	**80.68**
11.31447	**1,435**	**0.64**	**81.32**
11.3266	**1,382**	**0.61**	**81.93**
11.33857	**1,372**	**0.61**	**82.53**
11.35041	**1,442**	**0.64**	**83.17**
11.3621	**1,184**	**0.52**	**83.70**
11.37366	**1,205**	**0.53**	**84.23**
11.38509	**1,197**	**0.53**	**84.76**
11.39639	**1,237**	**0.55**	**85.31**
11.40757	**1,513**	**0.67**	**85.98**
11.41861	**1,207**	**0.53**	**86.51**
11.42954	**1,061**	**0.47**	**86.98**

11.44035	**949**	**0.42**	**87.40**
11.45105	**897**	**0.40**	**87.80**
11.46163	**893**	**0.40**	**88.19**
11.4721	**838**	**0.37**	**88.57**
11.48247	**754**	**0.33**	**88.90**
11.49272	**737**	**0.33**	**89.23**
11.50288	**728**	**0.32**	**89.55**
11.51293	**3,688**	**1.63**	**91.18**
11.60824	**4,638**	**2.05**	**93.23**
11.69525	**3,146**	**1.39**	**94.63**
11.71708	**166**	**0.07**	**94.70**
11.72276	**47**	**0.02**	**94.72**
11.76254	**73**	**0.03**	**94.75**
11.77529	**2,055**	**0.91**	**95.66**
11.77728	**24**	**0.01**	**95.67**
11.79492	**34**	**0.02**	**95.69**
11.79944	**83**	**0.04**	**95.72**
11.80117	**50**	**0.02**	**95.75**
11.8494	**1,417**	**0.63**	**96.37**
11.85418	**9**	**0.00**	**96.38**
11.89145	**30**	**0.01**	**96.39**
11.89441	**32**	**0.01**	**96.40**
11.90808	**198**	**0.09**	**96.49**
11.912	**14**	**0.01**	**96.50**
11.91484	**24**	**0.01**	**96.51**
11.91666	**36**	**0.02**	**96.53**
11.91695	**60**	**0.03**	**96.55**
11.91839	**1,100**	**0.49**	**97.04**
11.95348	**23**	**0.01**	**97.05**
11.96791	**32**	**0.01**	**97.06**
11.97572	**27**	**0.01**	**97.07**
11.98293	**754**	**0.33**	**97.41**
11.99546	**41**	**0.02**	**97.43**
12.0142	**38**	**0.02**	**97.44**
12.01851	**19**	**0.01**	**97.45**
12.03788	**39**	**0.02**	**97.47**
12.04355	**577**	**0.26**	**97.72**
12.075	**27**	**0.01**	**97.74**
12.10071	**373**	**0.17**	**97.90**
12.10606	**13**	**0.01**	**97.91**
12.12054	**83**	**0.04**	**97.94**
12.14043	**40**	**0.02**	**97.96**
12.15478	**291**	**0.13**	**98.09**
12.1626	**227**	**0.10**	**98.19**
12.1654	**26**	**0.01**	**98.20**
12.17861	**98**	**0.04**	**98.25**
12.18118	**161**	**0.07**	**98.32**
12.20607	**257**	**0.11**	**98.43**
12.20733	**59**	**0.03**	**98.46**
12.20761	**161**	**0.07**	**98.53**
12.24054	**8**	**0.00**	**98.53**
12.25486	**200**	**0.09**	**98.62**
12.26148	**31**	**0.01**	**98.63**
12.30138	**152**	**0.07**	**98.70**
12.30697	**25**	**0.01**	**98.71**
12.3099	**95**	**0.04**	**98.75**
12.34583	**117**	**0.05**	**98.81**
12.35372	**29**	**0.01**	**98.82**
12.37824	**17**	**0.01**	**98.83**
12.38839	**76**	**0.03**	**98.86**
12.4082	**7**	**0.00**	**98.86**
12.41875	**29**	**0.01**	**98.88**
12.42766	**30**	**0.01**	**98.89**
12.42922	**79**	**0.03**	**98.92**
12.46844	**55**	**0.02**	**98.95**

12.48217	38	0.02	98.97
12.49019	17	0.01	98.97
12.50618	34	0.02	98.99
12.51624	25	0.01	99.00
12.54255	41	0.02	99.02
12.54695	26	0.01	99.03
12.56228	30	0.01	99.04
12.57764	31	0.01	99.06
12.59368	376	0.17	99.22
12.59487	19	0.01	99.23
12.60426	195	0.09	99.32
12.61154	33	0.01	99.33
12.61602	70	0.03	99.36
12.61761	22	0.01	99.37
12.62881	86	0.04	99.41
12.64433	25	0.01	99.42
12.64489	18	0.01	99.43
12.64625	19	0.01	99.44
12.6532	15	0.01	99.44
12.67608	21	0.01	99.45
12.70685	17	0.01	99.46
12.7367	7	0.00	99.46
12.76569	5	0.00	99.47
12.77153	48	0.02	99.49
12.79386	5	0.00	99.49
12.80913	34	0.02	99.51
12.81843	30	0.01	99.52
12.82126	4	0.00	99.52
12.84482	32	0.01	99.53
12.84793	2	0.00	99.54
12.85931	16	0.01	99.54
12.8739	7	0.00	99.55
12.89922	4	0.00	99.55
12.92391	3	0.00	99.55
12.93541	204	0.09	99.64
12.94801	4	0.00	99.64
12.97154	3	0.00	99.64
12.97816	61	0.03	99.67
13.05472	31	0.01	99.68
13.08122	45	0.02	99.70
13.08236	32	0.01	99.72
13.08933	70	0.03	99.75
13.17992	66	0.03	99.78
13.2689	134	0.06	99.84
13.5294	301	0.13	99.97
13.88727	68	0.03	100.00
Total	225,912	100.00	

```
36 .
37 .
38 .
39 . *Howerver you will notice that if I use full time full year (49-52 weeks), I will have  far mor
```

```
40 .
41 . *defining a variable for gender
42 . gen MALE=0

43 . replace MALE=1 if(SEX==2)
(122,764 real changes made)

44 . label variable MALE "dummy variable if person's  Male"

45 . label define male 0 "Female" 1 "Male"

46 . label values MALE male

47 . tab MALE
```

dummy variable if person's Male	Freq.	Percent	Cum.
Female	**103,148**	**45.66**	**45.66**
Male	**122,764**	**54.34**	**100.00**
Total	**225,912**	**100.00**	

```
48 .
49 .
50 . *dropping data which is not available for minority
51 . drop if VISMIN==14
(1,066 observations deleted)

52 .
53 .
54 . *Dummy variable for person being a visible minority
55 .
56 . gen ALLVMIN=0

57 . replace ALLVMIN=1 if(VISMIN!=13)
(38,232 real changes made)

58 . label variable ALLVMIN "Dummy variable if a person is visible minority or not"

59 . label define allvmin 0 "White person" 1 "Visible minority"

60 . label values ALLVMIN allvmin

61 . tab ALLVMIN
```

Dummy variable if a person is visible minority or not	Freq.	Percent	Cum.
White person	**186,614**	**83.00**	**83.00**
Visible minority	**38,232**	**17.00**	**100.00**
Total	**224,846**	**100.00**	

```
62 .
63 .
64 . *defining South Asian (SA)  minority group variable
65 . gen S_A=0

66 . replace S_A=1 if (VISMIN==1)
(9,506 real changes made)

67 . label variable S_A "dummy variable if person is South Asian or not"

68 . label define s_a 0 "Not South Asian" 1 "South Asian"

69 . label values S_A  s_a

70 . tab S_A
```

dummy variable if person is South Asian or not	Freq.	Percent	Cum.
Not South Asian	**215,340**	**95.77**	**95.77**
South Asian	**9,506**	**4.23**	**100.00**
Total	**224,846**	**100.00**	

```
71 .
72 .
73 . *defining Chinese (CHI) minority group variable
74 .
75 . gen CHI=0

76 . replace CHI=1 if (VISMIN==2)
(8,802 real changes made)

77 . label variable CHI "dummy variable if person is Chinese or not"

78 . label define c_a 0 "Not chinese" 1 " Chinese"

79 . label values CHI  c_a

80 . tab CHI
```

dummy variable if person is Chinese or not	Freq.	Percent	Cum.
Not chinese	**216,044**	**96.09**	**96.09**
Chinese	**8,802**	**3.91**	**100.00**
Total	**224,846**	**100.00**	

```
81 .
82 . *defining Black (BL) minority group variable
83 . gen BL=0

84 . replace BL=1 if(VISMIN==3)
(5,221 real changes made)

85 . label variable BL "dummy variable if person is Black or not"

86 . label define bl 0 "Not Black" 1 " Black"

87 . label values BL  bl

88 . tab BL
```

dummy variable if person is Black or not	Freq.	Percent	Cum.
Not Black	**219,625**	**97.68**	**97.68**
Black	**5,221**	**2.32**	**100.00**
Total	**224,846**	**100.00**	

```
89 .
90 . *defining Flipino (FLP) minority group variable
91 . gen FLP=0

92 . replace FLP=1 if (VISMIN==4)
(5,043 real changes made)

93 . label variable S_A "dummy variable if person is Flipino or not"

94 . label define flp 0 "Not Filipino" 1 "Filipino"

95 . label values  FLP flp

96 . tab FLP
```

FLP	Freq.	Percent	Cum.
Not Filipino	**219,803**	**97.76**	**97.76**
Filipino	**5,043**	**2.24**	**100.00**
Total	**224,846**	**100.00**	

```
97 . *defining Latin American (LA) minority group variable
98 . gen LA=0

99 . replace LA=1 if (VISMIN==5)
(2,559 real changes made)
```

```
100 . label variable LA "dummy variable if person is Latin American or not"

101 . label define la 0 "Not Latin American" 1 " Latin American"

102 . label values LA  la

103 . tab LA
```

dummy variable if person is Latin American or not	Freq.	Percent	Cum.
Not Latin American	**222,287**	**98.86**	**98.86**
Latin American	**2,559**	**1.14**	**100.00**
Total	**224,846**	**100.00**	

```
104 .
105 . *defining Arab(Arb) minority group variable
106 . gen ARB=0

107 . replace ARB=1 if(VISMIN==6)
    (1,614 real changes made)

108 . label variable ARB "dummy variable if person is Arab or not"

109 . label define arb 0 "Not Arab" 1 "Arab"

110 . label values ARB  arb

111 . tab ARB
```

dummy variable if person is Arab or not	Freq.	Percent	Cum.
Not Arab	**223,232**	**99.28**	**99.28**
Arab	**1,614**	**0.72**	**100.00**
Total	**224,846**	**100.00**	

```
112 . *defining Southeast Asian (SE_A) minority group variable
113 .
114 .
115 . gen SE_A=0

116 . replace SE_A=1 if (VISMIN==7)
    (2,053 real changes made)

117 . label variable SE_A "dummy variable if person is Southeast Asian or not"

118 . label define se_a 0 "Not Southeast Asian" 1 " Southeast Asian"
```

```
119 . label values SE_A  se_a

120 . tab SE_A
```

dummy variable if person is Southeast Asian or not	Freq.	Percent	Cum.
Not Southeast Asian	**222,793**	**99.09**	**99.09**
Southeast Asian	**2,053**	**0.91**	**100.00**
Total	**224,846**	**100.00**	

```
121 . *defining West Asian (W_A) minority group variable
122 .
123 . gen W_A=0

124 . replace W_A=1 if (VISMIN==8)
(953 real changes made)

125 . label variable W_A "dummy variable if person is West Asian or not"

126 . label define w_a 0 "Not West Asian" 1 "West Asian"

127 . label values W_A  w_a

128 . tab W_A
```

dummy variable if person is West Asian or not	Freq.	Percent	Cum.
Not West Asian	**223,893**	**99.58**	**99.58**
West Asian	**953**	**0.42**	**100.00**
Total	**224,846**	**100.00**	

```
129 .
130 .
131 . *defining Korean(KO) minority group variable
132 .
133 . gen KO=0

134 . replace KO=1 if (VISMIN==9)
(590 real changes made)

135 . label variable KO "dummy variable if person is Korean or not"

136 . label define ko 0 "Not Korean" 1 "Korean"

137 . label values KO  ko
```

```
138 . tab KO
```

dummy variable if person is Korean or not	Freq.	Percent	Cum.
Not Korean	**224,256**	**99.74**	**99.74**
Korean	**590**	**0.26**	**100.00**
Total	**224,846**	**100.00**	

```
139 .
140 . *defining Japanese(JP) minority group variable
141 .
142 . gen JP=0

143 . replace JP=1 if (VISMIN==10)
(441 real changes made)

144 . label variable JP "dummy variable if person is Japanese or not"

145 . label define jp 0 "Not Japanese" 1 "Japanese"

146 . label values JP  jp

147 .
148 . tab JP
```

dummy variable if person is Japanese or not	Freq.	Percent	Cum.
Not Japanese	**224,405**	**99.80**	**99.80**
Japanese	**441**	**0.20**	**100.00**
Total	**224,846**	**100.00**	

```
149 .
150 . *defining Variable for Visible minorities not included elsewhere
151 .
152 . gen VMN=0

153 . replace VMN=1 if (VISMIN==11)
(643 real changes made)

154 . label variable VMN "dummy variable if person is Visible minorities not included elsewhere"

155 . label define vmn 1 "VM Not included elsewhere " 0 "Not VM or included elsewhere"
```

```
156 . label values VMN vmn

157 . tab VMN
```

dummy variable if person is Visible minorities not included elsewhere	Freq.	Percent	Cum.
Not VM or included elsewhere	**224,203**	**99.71**	**99.71**
VM Not included elsewhere	**643**	**0.29**	**100.00**
Total	**224,846**	**100.00**	

```
158 .
159 . *defining Variable for  Aboriginal minority
160 .
161 .  gen AB=0

162 . replace AB=1 if (ETHDER==1)
(2,528 real changes made)

163 . label variable AB "dummy variable if person is Aboriginal or not"

164 . label define ab 0 "Not Aborginal" 1 " Aboriginal"

165 . label values AB ab

166 . tab AB
```

dummy variable if person is Aboriginal or not	Freq.	Percent	Cum.
Not Aborginal	**222,318**	**98.88**	**98.88**
Aboriginal	**2,528**	**1.12**	**100.00**
Total	**224,846**	**100.00**	

```
167 .
168 .
169 . *defining variable for multiple visible minorities
170 . gen MVM=0

171 . replace MVM=1 if (VISMIN==12)
(807 real changes made)

172 . label variable MVM "dummy variable if person is multiple visible minorities or not"

173 . label define mvm 0 "Not Multiple VM" 1 "Multiple VM"

174 . label values MVM mvm
```

```
175 . tab MVM
```

dummy variable if person is multiple visible minorities or not	Freq.	Percent	Cum.
Not Multiple VM	**224,039**	**99.64**	**99.64**
Multiple VM	**807**	**0.36**	**100.00**
Total	**224,846**	**100.00**	

```
176 .
177 .
178 .
179 . *defining a variable for Apprenticeship certificate and diploma level education
180 .  gen ADEDUC=0

181 . replace ADEDUC=1 if(HDGREE==3|HDGREE==4)
(25,728 real changes made)

182 . label variable ADEDUC "dummy variable if person has diploma, apprenticeship or not"

183 . label define dued 0 "No diploma nor Apprenticeship" 1 "Apprenticeship"

184 . label values ADEDUC dued

185 . tab ADEDUC
```

dummy variable if person has diploma, apprenticeship or not	Freq.	Percent	Cum.
No diploma nor Apprenticeship	**199,118**	**88.56**	**88.56**
Apprenticeship	**25,728**	**11.44**	**100.00**
Total	**224,846**	**100.00**	

```
186 .
187 . *defining variable for college and university level education
188 . gen UEDUC=0

189 . replace UEDUC=1 if(inrange(HDGREE,5,10))
(112,818 real changes made)

190 . label variable UEDUC "dummy variable if person has Undergraduate degree or diploma above underg
note: label truncated to 80 characters

191 . label define ued 0 "No University/college Education" 1 "University college Education"

192 . label values UEDUC ued
```

```
193 . tab UEDUC
```

dummy variable if person has Undergraduate degree or diploma above undergraduate	Freq.	Percent	Cum.
No University/college Education	**112,028**	**49.82**	**49.82**
University college Education	**112,818**	**50.18**	**100.00**
Total	**224,846**	**100.00**	

```
194 .
195 . *defining variable for which education achieved is Highschool diploma or equivalent
196 .
197 . gen HEDUC=0

198 . replace HEDUC=1 if(HDGREE==2)
(50,611 real changes made)

199 . label variable HEDUC "dummy variable if person has highschool education or not"

200 . label define hed 1 "Educ is Highschool" 0 "Educ not Highschool"

201 . label values HEDUC hed

202 . tab HEDUC
```

dummy variable if person has highschool education or not	Freq.	Percent	Cum.
Educ not Highschool	**174,235**	**77.49**	**77.49**
Educ is Highschool	**50,611**	**22.51**	**100.00**
Total	**224,846**	**100.00**	

```
203 .
204 .  *defining a variable for Masters and Doctrate level Education
205 .  gen DEDUC=0

206 .  replace DEDUC=1 if(inrange(HDGREE,11,13))
(15,978 real changes made)

207 .  label variable DEDUC "dummy variable if person has masters or doctrate"

208 .  label define deduc 0 "No doctrate or masters"  1 "Doctrate or Masters"

209 .  label values DEDUC deduc

210 .  tab DEDUC
```

dummy variable if person has masters or doctrate	Freq.	Percent	Cum.
No doctrate or masters	**208,868**	**92.89**	**92.89**
Doctrate or Masters	**15,978**	**7.11**	**100.00**
Total	**224,846**	**100.00**	

```
211 .
212 .  *I have grouped Education level into 4 groups, I have kept people with masters and doctrates i
    > rtificate in one.
213 .
214 .  *Dropping data for when data is not avalaible
215 .  drop if HDGREE==14
    (1,231 observations deleted)

216 .
217 .
218 .
    end of do-file

219 .
```