8-FINAL

November 26, 2018

1 Assignment 8

Using any dataset that has a categorical feature that needs to be predicted, use several algorithms, preprocessing techniques, feature extraction techniques to fit the data to the model and show the accuracy, confusion matrix, and the classification report. G https://www.kaggle.com/ntnu-testimon/paysim1 https://www.kaggle.com/joniarroba/noshowappointments https://archive.ics.uci.edu/ml/datasets.html?format=&task=cla&att=&area=&numAtt=&numIns=&type=&sort https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients https://archive.ics.uci.edu/ml/datasets/Adult

```
In [12]: from sklearn.svm import LinearSVC
         from sklearn.linear_model import LogisticRegression
         import pandas as pd
         import seaborn as sns
         import matplotlib.pyplot as plt
         %matplotlib inline
         sns.set(font_scale=1.5)
         import numpy as np

         from pylab import rcParams
         rcParams['figure.figsize'] = 20, 10


         from sklearn.linear_model import LogisticRegression as Model
```

```
In [13]: # Read in the Adult data [https://archive.ics.uci.edu/ml/datasets/Adult]
```

```
In [15]: adult = pd.read_csv('../data/adult_data.csv')
         adult
```

```
Out[15]:      age            workclass  fnlwgt   education  education_num  \
         0     39            State-gov   77516   Bachelors             13
         1     50     Self-emp-not-inc   83311   Bachelors             13
         2     38              Private  215646     HS-grad              9
         3     53              Private  234721        11th              7
         4     28              Private  338409   Bachelors             13
         5     37              Private  284582     Masters             14
         6     49              Private  160187         9th              5
```

1

7	52	Self-emp-not-inc	209642	HS-grad	9
8	31	Private	45781	Masters	14
9	42	Private	159449	Bachelors	13
10	37	Private	280464	Some-college	10
11	30	State-gov	141297	Bachelors	13
12	23	Private	122272	Bachelors	13
13	32	Private	205019	Assoc-acdm	12
14	40	Private	121772	Assoc-voc	11
15	34	Private	245487	7th-8th	4
16	25	Self-emp-not-inc	176756	HS-grad	9
17	32	Private	186824	HS-grad	9
18	38	Private	28887	11th	7
19	43	Self-emp-not-inc	292175	Masters	14
20	40	Private	193524	Doctorate	16
21	54	Private	302146	HS-grad	9
22	35	Federal-gov	76845	9th	5
23	43	Private	117037	11th	7
24	59	Private	109015	HS-grad	9
25	56	Local-gov	216851	Bachelors	13
26	19	Private	168294	HS-grad	9
27	54	?	180211	Some-college	10
28	39	Private	367260	HS-grad	9
29	49	Private	193366	HS-grad	9
...	...	...	...	...	...
32531	30	?	33811	Bachelors	13
32532	34	Private	204461	Doctorate	16
32533	54	Private	337992	Bachelors	13
32534	37	Private	179137	Some-college	10
32535	22	Private	325033	12th	8
32536	34	Private	160216	Bachelors	13
32537	30	Private	345898	HS-grad	9
32538	38	Private	139180	Bachelors	13
32539	71	?	287372	Doctorate	16
32540	45	State-gov	252208	HS-grad	9
32541	41	?	202822	HS-grad	9
32542	72	?	129912	HS-grad	9
32543	45	Local-gov	119199	Assoc-acdm	12
32544	31	Private	199655	Masters	14
32545	39	Local-gov	111499	Assoc-acdm	12
32546	37	Private	198216	Assoc-acdm	12
32547	43	Private	260761	HS-grad	9
32548	65	Self-emp-not-inc	99359	Prof-school	15
32549	43	State-gov	255835	Some-college	10
32550	43	Self-emp-not-inc	27242	Some-college	10
32551	32	Private	34066	10th	6
32552	43	Private	84661	Assoc-voc	11
32553	32	Private	116138	Masters	14
32554	53	Private	321865	Masters	14

```
32555  22          Private  310152    Some-college           10
32556  27          Private  257302     Assoc-acdm           12
32557  40          Private  154374        HS-grad            9
32558  58          Private  151910        HS-grad            9
32559  22          Private  201490        HS-grad            9
32560  52     Self-emp-inc  287927        HS-grad            9
```

	marital_status	occupation	relationship
0	Never-married	Adm-clerical	Not-in-family
1	Married-civ-spouse	Exec-managerial	Husband
2	Divorced	Handlers-cleaners	Not-in-family
3	Married-civ-spouse	Handlers-cleaners	Husband
4	Married-civ-spouse	Prof-specialty	Wife
5	Married-civ-spouse	Exec-managerial	Wife
6	Married-spouse-absent	Other-service	Not-in-family
7	Married-civ-spouse	Exec-managerial	Husband
8	Never-married	Prof-specialty	Not-in-family
9	Married-civ-spouse	Exec-managerial	Husband
10	Married-civ-spouse	Exec-managerial	Husband
11	Married-civ-spouse	Prof-specialty	Husband
12	Never-married	Adm-clerical	Own-child
13	Never-married	Sales	Not-in-family
14	Married-civ-spouse	Craft-repair	Husband
15	Married-civ-spouse	Transport-moving	Husband
16	Never-married	Farming-fishing	Own-child
17	Never-married	Machine-op-inspct	Unmarried
18	Married-civ-spouse	Sales	Husband
19	Divorced	Exec-managerial	Unmarried
20	Married-civ-spouse	Prof-specialty	Husband
21	Separated	Other-service	Unmarried
22	Married-civ-spouse	Farming-fishing	Husband
23	Married-civ-spouse	Transport-moving	Husband
24	Divorced	Tech-support	Unmarried
25	Married-civ-spouse	Tech-support	Husband
26	Never-married	Craft-repair	Own-child
27	Married-civ-spouse	?	Husband
28	Divorced	Exec-managerial	Not-in-family
29	Married-civ-spouse	Craft-repair	Husband
...	...	...	...
32531	Never-married	?	Not-in-family
32532	Married-civ-spouse	Prof-specialty	Husband
32533	Married-civ-spouse	Exec-managerial	Husband
32534	Divorced	Adm-clerical	Unmarried
32535	Never-married	Protective-serv	Own-child
32536	Never-married	Exec-managerial	Not-in-family
32537	Never-married	Craft-repair	Not-in-family
32538	Divorced	Prof-specialty	Unmarried
32539	Married-civ-spouse	?	Husband

```
32540                Separated       Adm-clerical        Own-child
32541                Separated                  ?     Not-in-family
32542        Married-civ-spouse                  ?          Husband
32543                 Divorced      Prof-specialty        Unmarried
32544                 Divorced       Other-service     Not-in-family
32545        Married-civ-spouse       Adm-clerical             Wife
32546                 Divorced        Tech-support     Not-in-family
32547        Married-civ-spouse   Machine-op-inspct          Husband
32548            Never-married      Prof-specialty     Not-in-family
32549                 Divorced        Adm-clerical    Other-relative
32550        Married-civ-spouse        Craft-repair          Husband
32551        Married-civ-spouse   Handlers-cleaners          Husband
32552        Married-civ-spouse               Sales          Husband
32553            Never-married        Tech-support     Not-in-family
32554        Married-civ-spouse     Exec-managerial          Husband
32555            Never-married     Protective-serv     Not-in-family
32556        Married-civ-spouse        Tech-support             Wife
32557        Married-civ-spouse   Machine-op-inspct          Husband
32558                  Widowed        Adm-clerical        Unmarried
32559            Never-married        Adm-clerical        Own-child
32560        Married-civ-spouse     Exec-managerial             Wife

                     race      sex   capital_gain   capital_loss   hpw  \
0                   White     Male           2174              0    40
1                   White     Male              0              0    13
2                   White     Male              0              0    40
3                   Black     Male              0              0    40
4                   Black   Female              0              0    40
5                   White   Female              0              0    40
6                   Black   Female              0              0    16
7                   White     Male              0              0    45
8                   White   Female          14084              0    50
9                   White     Male           5178              0    40
10                  Black     Male              0              0    80
11     Asian-Pac-Islander     Male              0              0    40
12                  White   Female              0              0    30
13                  Black     Male              0              0    50
14     Asian-Pac-Islander     Male              0              0    40
15     Amer-Indian-Eskimo     Male              0              0    45
16                  White     Male              0              0    35
17                  White     Male              0              0    40
18                  White     Male              0              0    50
19                  White   Female              0              0    45
20                  White     Male              0              0    60
21                  Black   Female              0              0    20
22                  Black     Male              0              0    40
23                  White     Male              0           2042    40
24                  White   Female              0              0    40
```

25	White	Male	0	0	40
26	White	Male	0	0	40
27	Asian-Pac-Islander	Male	0	0	60
28	White	Male	0	0	80
29	White	Male	0	0	40
...	...	...	...	...	...
32531	Asian-Pac-Islander	Female	0	0	99
32532	White	Male	0	0	60
32533	Asian-Pac-Islander	Male	0	0	50
32534	White	Female	0	0	39
32535	Black	Male	0	0	35
32536	White	Female	0	0	55
32537	Black	Male	0	0	46
32538	Black	Female	15020	0	45
32539	White	Male	0	0	10
32540	White	Female	0	0	40
32541	Black	Female	0	0	32
32542	White	Male	0	0	25
32543	White	Female	0	0	48
32544	Other	Female	0	0	30
32545	White	Female	0	0	20
32546	White	Female	0	0	40
32547	White	Male	0	0	40
32548	White	Male	1086	0	60
32549	White	Female	0	0	40
32550	White	Male	0	0	50
32551	Amer-Indian-Eskimo	Male	0	0	40
32552	White	Male	0	0	45
32553	Asian-Pac-Islander	Male	0	0	11
32554	White	Male	0	0	40
32555	White	Male	0	0	40
32556	White	Female	0	0	38
32557	White	Male	0	0	40
32558	White	Female	0	0	40
32559	White	Male	0	0	20
32560	White	Female	15024	0	40

	native_country	salary_50K
0	United-States	<=50K
1	United-States	<=50K
2	United-States	<=50K
3	United-States	<=50K
4	Cuba	<=50K
5	United-States	<=50K
6	Jamaica	<=50K
7	United-States	>50K
8	United-States	>50K
9	United-States	>50K

```
10      United-States      >50K
11              India       >50K
12      United-States     <=50K
13      United-States     <=50K
14                  ?      >50K
15             Mexico     <=50K
16      United-States     <=50K
17      United-States     <=50K
18      United-States     <=50K
19      United-States      >50K
20      United-States      >50K
21      United-States     <=50K
22      United-States     <=50K
23      United-States     <=50K
24      United-States     <=50K
25      United-States      >50K
26      United-States     <=50K
27              South       >50K
28      United-States     <=50K
29      United-States     <=50K
...                ...       ...
32531   United-States     <=50K
32532   United-States      >50K
32533           Japan       >50K
32534   United-States     <=50K
32535   United-States     <=50K
32536   United-States      >50K
32537   United-States     <=50K
32538   United-States      >50K
32539   United-States      >50K
32540   United-States     <=50K
32541   United-States     <=50K
32542   United-States     <=50K
32543   United-States     <=50K
32544   United-States     <=50K
32545   United-States      >50K
32546   United-States     <=50K
32547          Mexico     <=50K
32548   United-States     <=50K
32549   United-States     <=50K
32550   United-States     <=50K
32551   United-States     <=50K
32552   United-States     <=50K
32553          Taiwan     <=50K
32554   United-States      >50K
32555   United-States     <=50K
32556   United-States     <=50K
32557   United-States      >50K
```

```
         32558    United-States        <=50K
         32559    United-States        <=50K
         32560    United-States         >50K

         [32561 rows x 15 columns]
```

In [16]: `from sklearn.svm import LinearSVC`

In [17]: `adult.columns`

Out[17]: `Index(['age', 'workclass', 'fnlwgt', 'education', 'education_num',`
`       'marital_status', 'occupation', 'relationship', 'race', 'sex',`
`       'capital_gain', 'capital_loss', 'hpw', 'native_country', 'salary_50K'],`
`      dtype='object')`

In [19]: `adult.groupby(adult.salary_50K).count()`

Out[19]:

	age	workclass	fnlwgt	education	education_num
salary_50K					
<=50K	24720	24720	24720	24720	24720
>50K	7841	7841	7841	7841	7841

	marital_status	occupation	relationship	race	sex
salary_50K					
<=50K	24720	24720	24720	24720	24720
>50K	7841	7841	7841	7841	7841

	capital_gain	capital_loss	hpw	native_country
salary_50K				
<=50K	24720	24720	24720	24720
>50K	7841	7841	7841	7841

In [20]: `adult.iloc[0]`

Out[20]:
```
age                             39
workclass                State-gov
fnlwgt                       77516
education                 Bachelors
education_num                   13
marital_status       Never-married
occupation            Adm-clerical
relationship        Not-in-family
race                         White
sex                           Male
capital_gain                  2174
capital_loss                     0
hpw                             40
native_country       United-States
salary_50K                   <=50K
Name: 0, dtype: object
```

```
In [22]: df = pd.DataFrame(data=adult)
         df

Out[22]:        age        workclass   fnlwgt     education  education_num  \
         0       39        State-gov    77516     Bachelors             13
         1       50  Self-emp-not-inc   83311     Bachelors             13
         2       38          Private   215646      HS-grad              9
         3       53          Private   234721         11th              7
         4       28          Private   338409     Bachelors             13
         5       37          Private   284582       Masters             14
         6       49          Private   160187          9th              5
         7       52  Self-emp-not-inc  209642      HS-grad              9
         8       31          Private    45781       Masters             14
         9       42          Private   159449     Bachelors             13
         10      37          Private   280464  Some-college             10
         11      30        State-gov   141297     Bachelors             13
         12      23          Private   122272     Bachelors             13
         13      32          Private   205019    Assoc-acdm             12
         14      40          Private   121772     Assoc-voc             11
         15      34          Private   245487      7th-8th              4
         16      25  Self-emp-not-inc  176756      HS-grad              9
         17      32          Private   186824      HS-grad              9
         18      38          Private    28887         11th              7
         19      43  Self-emp-not-inc  292175       Masters             14
         20      40          Private   193524     Doctorate            16
         21      54          Private   302146      HS-grad              9
         22      35       Federal-gov   76845          9th              5
         23      43          Private   117037         11th              7
         24      59          Private   109015      HS-grad              9
         25      56        Local-gov   216851     Bachelors             13
         26      19          Private   168294      HS-grad              9
         27      54                ?   180211  Some-college             10
         28      39          Private   367260      HS-grad              9
         29      49          Private   193366      HS-grad              9
         ...    ...              ...      ...           ...            ...
         32531   30                ?    33811     Bachelors             13
         32532   34          Private   204461     Doctorate            16
         32533   54          Private   337992     Bachelors             13
         32534   37          Private   179137  Some-college             10
         32535   22          Private   325033         12th              8
         32536   34          Private   160216     Bachelors             13
         32537   30          Private   345898      HS-grad              9
         32538   38          Private   139180     Bachelors             13
         32539   71                ?   287372     Doctorate            16
         32540   45        State-gov   252208      HS-grad              9
         32541   41                ?   202822      HS-grad              9
         32542   72                ?   129912      HS-grad              9
         32543   45        Local-gov   119199    Assoc-acdm             12
```

```
32544  31            Private  199655       Masters          14
32545  39          Local-gov  111499     Assoc-acdm          12
32546  37            Private  198216     Assoc-acdm          12
32547  43            Private  260761        HS-grad           9
32548  65    Self-emp-not-inc   99359    Prof-school          15
32549  43          State-gov  255835   Some-college          10
32550  43    Self-emp-not-inc   27242   Some-college          10
32551  32            Private   34066           10th           6
32552  43            Private   84661      Assoc-voc          11
32553  32            Private  116138        Masters          14
32554  53            Private  321865        Masters          14
32555  22            Private  310152   Some-college          10
32556  27            Private  257302     Assoc-acdm          12
32557  40            Private  154374        HS-grad           9
32558  58            Private  151910        HS-grad           9
32559  22            Private  201490        HS-grad           9
32560  52        Self-emp-inc  287927        HS-grad           9

            marital_status          occupation   relationship  \
0            Never-married        Adm-clerical   Not-in-family
1       Married-civ-spouse     Exec-managerial        Husband
2                 Divorced   Handlers-cleaners   Not-in-family
3       Married-civ-spouse   Handlers-cleaners        Husband
4       Married-civ-spouse      Prof-specialty           Wife
5       Married-civ-spouse     Exec-managerial           Wife
6     Married-spouse-absent       Other-service   Not-in-family
7       Married-civ-spouse     Exec-managerial        Husband
8            Never-married      Prof-specialty   Not-in-family
9       Married-civ-spouse     Exec-managerial        Husband
10      Married-civ-spouse     Exec-managerial        Husband
11      Married-civ-spouse      Prof-specialty        Husband
12           Never-married        Adm-clerical      Own-child
13           Never-married               Sales   Not-in-family
14      Married-civ-spouse        Craft-repair        Husband
15      Married-civ-spouse     Transport-moving        Husband
16           Never-married      Farming-fishing      Own-child
17           Never-married   Machine-op-inspct      Unmarried
18      Married-civ-spouse               Sales        Husband
19                Divorced     Exec-managerial      Unmarried
20      Married-civ-spouse      Prof-specialty        Husband
21               Separated       Other-service      Unmarried
22      Married-civ-spouse     Farming-fishing        Husband
23      Married-civ-spouse     Transport-moving        Husband
24                Divorced        Tech-support      Unmarried
25      Married-civ-spouse        Tech-support        Husband
26           Never-married        Craft-repair      Own-child
27      Married-civ-spouse                   ?        Husband
28                Divorced     Exec-managerial   Not-in-family
```

```
29          Married-civ-spouse        Craft-repair           Husband
...                ...                    ...                    ...
32531          Never-married                  ?        Not-in-family
32532      Married-civ-spouse      Prof-specialty           Husband
32533      Married-civ-spouse     Exec-managerial           Husband
32534            Divorced          Adm-clerical           Unmarried
32535          Never-married      Protective-serv           Own-child
32536          Never-married      Exec-managerial       Not-in-family
32537          Never-married         Craft-repair       Not-in-family
32538            Divorced         Prof-specialty           Unmarried
32539      Married-civ-spouse                 ?             Husband
32540            Separated          Adm-clerical           Own-child
32541            Separated                  ?         Not-in-family
32542      Married-civ-spouse                 ?             Husband
32543            Divorced         Prof-specialty           Unmarried
32544            Divorced          Other-service       Not-in-family
32545      Married-civ-spouse       Adm-clerical                Wife
32546            Divorced          Tech-support        Not-in-family
32547      Married-civ-spouse    Machine-op-inspct           Husband
32548          Never-married      Prof-specialty        Not-in-family
32549            Divorced          Adm-clerical       Other-relative
32550      Married-civ-spouse        Craft-repair           Husband
32551      Married-civ-spouse    Handlers-cleaners           Husband
32552      Married-civ-spouse              Sales            Husband
32553          Never-married        Tech-support        Not-in-family
32554      Married-civ-spouse     Exec-managerial           Husband
32555          Never-married      Protective-serv       Not-in-family
32556      Married-civ-spouse        Tech-support                Wife
32557      Married-civ-spouse    Machine-op-inspct           Husband
32558            Widowed           Adm-clerical           Unmarried
32559          Never-married        Adm-clerical           Own-child
32560      Married-civ-spouse     Exec-managerial                Wife

                    race     sex  capital_gain  capital_loss  hpw  \
0                  White    Male          2174             0   40
1                  White    Male             0             0   13
2                  White    Male             0             0   40
3                  Black    Male             0             0   40
4                  Black  Female             0             0   40
5                  White  Female             0             0   40
6                  Black  Female             0             0   16
7                  White    Male             0             0   45
8                  White  Female         14084             0   50
9                  White    Male          5178             0   40
10                 Black    Male             0             0   80
11    Asian-Pac-Islander    Male             0             0   40
12                 White  Female             0             0   30
13                 Black    Male             0             0   50
```

14	Asian-Pac-Islander	Male	0	0	40
15	Amer-Indian-Eskimo	Male	0	0	45
16	White	Male	0	0	35
17	White	Male	0	0	40
18	White	Male	0	0	50
19	White	Female	0	0	45
20	White	Male	0	0	60
21	Black	Female	0	0	20
22	Black	Male	0	0	40
23	White	Male	0	2042	40
24	White	Female	0	0	40
25	White	Male	0	0	40
26	White	Male	0	0	40
27	Asian-Pac-Islander	Male	0	0	60
28	White	Male	0	0	80
29	White	Male	0	0	40
...	...	...	...	...	...
32531	Asian-Pac-Islander	Female	0	0	99
32532	White	Male	0	0	60
32533	Asian-Pac-Islander	Male	0	0	50
32534	White	Female	0	0	39
32535	Black	Male	0	0	35
32536	White	Female	0	0	55
32537	Black	Male	0	0	46
32538	Black	Female	15020	0	45
32539	White	Male	0	0	10
32540	White	Female	0	0	40
32541	Black	Female	0	0	32
32542	White	Male	0	0	25
32543	White	Female	0	0	48
32544	Other	Female	0	0	30
32545	White	Female	0	0	20
32546	White	Female	0	0	40
32547	White	Male	0	0	40
32548	White	Male	1086	0	60
32549	White	Female	0	0	40
32550	White	Male	0	0	50
32551	Amer-Indian-Eskimo	Male	0	0	40
32552	White	Male	0	0	45
32553	Asian-Pac-Islander	Male	0	0	11
32554	White	Male	0	0	40
32555	White	Male	0	0	40
32556	White	Female	0	0	38
32557	White	Male	0	0	40
32558	White	Female	0	0	40
32559	White	Male	0	0	20
32560	White	Female	15024	0	40

```
       native_country  salary_50K
0        United-States       <=50K
1        United-States       <=50K
2        United-States       <=50K
3        United-States       <=50K
4                 Cuba       <=50K
5        United-States       <=50K
6              Jamaica       <=50K
7        United-States        >50K
8        United-States        >50K
9        United-States        >50K
10       United-States        >50K
11               India        >50K
12       United-States       <=50K
13       United-States       <=50K
14                   ?        >50K
15              Mexico       <=50K
16       United-States       <=50K
17       United-States       <=50K
18       United-States       <=50K
19       United-States        >50K
20       United-States        >50K
21       United-States       <=50K
22       United-States       <=50K
23       United-States       <=50K
24       United-States       <=50K
25       United-States        >50K
26       United-States       <=50K
27               South        >50K
28       United-States       <=50K
29       United-States       <=50K
...                ...         ...
32531    United-States       <=50K
32532    United-States        >50K
32533            Japan        >50K
32534    United-States       <=50K
32535    United-States       <=50K
32536    United-States        >50K
32537    United-States       <=50K
32538    United-States        >50K
32539    United-States        >50K
32540    United-States       <=50K
32541    United-States       <=50K
32542    United-States       <=50K
32543    United-States       <=50K
32544    United-States       <=50K
32545    United-States        >50K
32546    United-States       <=50K
```

```
32547           Mexico    <=50K
32548    United-States    <=50K
32549    United-States    <=50K
32550    United-States    <=50K
32551    United-States    <=50K
32552    United-States    <=50K
32553           Taiwan    <=50K
32554    United-States     >50K
32555    United-States    <=50K
32556    United-States    <=50K
32557    United-States     >50K
32558    United-States    <=50K
32559    United-States    <=50K
32560    United-States     >50K

[32561 rows x 15 columns]

In [23]: df['workclass_num']= df['workclass']
         df

Out[23]:      age          workclass  fnlwgt     education  education_num  \
         0     39          State-gov   77516     Bachelors             13
         1     50   Self-emp-not-inc   83311     Bachelors             13
         2     38            Private  215646       HS-grad              9
         3     53            Private  234721          11th              7
         4     28            Private  338409     Bachelors             13
         5     37            Private  284582       Masters             14
         6     49            Private  160187           9th              5
         7     52   Self-emp-not-inc  209642       HS-grad              9
         8     31            Private   45781       Masters             14
         9     42            Private  159449     Bachelors             13
         10    37            Private  280464  Some-college             10
         11    30          State-gov  141297     Bachelors             13
         12    23            Private  122272     Bachelors             13
         13    32            Private  205019     Assoc-acdm             12
         14    40            Private  121772      Assoc-voc             11
         15    34            Private  245487       7th-8th              4
         16    25   Self-emp-not-inc  176756       HS-grad              9
         17    32            Private  186824       HS-grad              9
         18    38            Private   28887          11th              7
         19    43   Self-emp-not-inc  292175       Masters             14
         20    40            Private  193524     Doctorate             16
         21    54            Private  302146       HS-grad              9
         22    35        Federal-gov   76845           9th              5
         23    43            Private  117037          11th              7
         24    59            Private  109015       HS-grad              9
         25    56          Local-gov  216851     Bachelors             13
         26    19            Private  168294       HS-grad              9
```

27	54	?	180211	Some-college	10
28	39	Private	367260	HS-grad	9
29	49	Private	193366	HS-grad	9
...	...	...	...	...	...
32531	30	?	33811	Bachelors	13
32532	34	Private	204461	Doctorate	16
32533	54	Private	337992	Bachelors	13
32534	37	Private	179137	Some-college	10
32535	22	Private	325033	12th	8
32536	34	Private	160216	Bachelors	13
32537	30	Private	345898	HS-grad	9
32538	38	Private	139180	Bachelors	13
32539	71	?	287372	Doctorate	16
32540	45	State-gov	252208	HS-grad	9
32541	41	?	202822	HS-grad	9
32542	72	?	129912	HS-grad	9
32543	45	Local-gov	119199	Assoc-acdm	12
32544	31	Private	199655	Masters	14
32545	39	Local-gov	111499	Assoc-acdm	12
32546	37	Private	198216	Assoc-acdm	12
32547	43	Private	260761	HS-grad	9
32548	65	Self-emp-not-inc	99359	Prof-school	15
32549	43	State-gov	255835	Some-college	10
32550	43	Self-emp-not-inc	27242	Some-college	10
32551	32	Private	34066	10th	6
32552	43	Private	84661	Assoc-voc	11
32553	32	Private	116138	Masters	14
32554	53	Private	321865	Masters	14
32555	22	Private	310152	Some-college	10
32556	27	Private	257302	Assoc-acdm	12
32557	40	Private	154374	HS-grad	9
32558	58	Private	151910	HS-grad	9
32559	22	Private	201490	HS-grad	9
32560	52	Self-emp-inc	287927	HS-grad	9

	marital_status	occupation	relationship \
0	Never-married	Adm-clerical	Not-in-family
1	Married-civ-spouse	Exec-managerial	Husband
2	Divorced	Handlers-cleaners	Not-in-family
3	Married-civ-spouse	Handlers-cleaners	Husband
4	Married-civ-spouse	Prof-specialty	Wife
5	Married-civ-spouse	Exec-managerial	Wife
6	Married-spouse-absent	Other-service	Not-in-family
7	Married-civ-spouse	Exec-managerial	Husband
8	Never-married	Prof-specialty	Not-in-family
9	Married-civ-spouse	Exec-managerial	Husband
10	Married-civ-spouse	Exec-managerial	Husband
11	Married-civ-spouse	Prof-specialty	Husband

12	Never-married	Adm-clerical	Own-child
13	Never-married	Sales	Not-in-family
14	Married-civ-spouse	Craft-repair	Husband
15	Married-civ-spouse	Transport-moving	Husband
16	Never-married	Farming-fishing	Own-child
17	Never-married	Machine-op-inspct	Unmarried
18	Married-civ-spouse	Sales	Husband
19	Divorced	Exec-managerial	Unmarried
20	Married-civ-spouse	Prof-specialty	Husband
21	Separated	Other-service	Unmarried
22	Married-civ-spouse	Farming-fishing	Husband
23	Married-civ-spouse	Transport-moving	Husband
24	Divorced	Tech-support	Unmarried
25	Married-civ-spouse	Tech-support	Husband
26	Never-married	Craft-repair	Own-child
27	Married-civ-spouse	?	Husband
28	Divorced	Exec-managerial	Not-in-family
29	Married-civ-spouse	Craft-repair	Husband
...	...	...	...
32531	Never-married	?	Not-in-family
32532	Married-civ-spouse	Prof-specialty	Husband
32533	Married-civ-spouse	Exec-managerial	Husband
32534	Divorced	Adm-clerical	Unmarried
32535	Never-married	Protective-serv	Own-child
32536	Never-married	Exec-managerial	Not-in-family
32537	Never-married	Craft-repair	Not-in-family
32538	Divorced	Prof-specialty	Unmarried
32539	Married-civ-spouse	?	Husband
32540	Separated	Adm-clerical	Own-child
32541	Separated	?	Not-in-family
32542	Married-civ-spouse	?	Husband
32543	Divorced	Prof-specialty	Unmarried
32544	Divorced	Other-service	Not-in-family
32545	Married-civ-spouse	Adm-clerical	Wife
32546	Divorced	Tech-support	Not-in-family
32547	Married-civ-spouse	Machine-op-inspct	Husband
32548	Never-married	Prof-specialty	Not-in-family
32549	Divorced	Adm-clerical	Other-relative
32550	Married-civ-spouse	Craft-repair	Husband
32551	Married-civ-spouse	Handlers-cleaners	Husband
32552	Married-civ-spouse	Sales	Husband
32553	Never-married	Tech-support	Not-in-family
32554	Married-civ-spouse	Exec-managerial	Husband
32555	Never-married	Protective-serv	Not-in-family
32556	Married-civ-spouse	Tech-support	Wife
32557	Married-civ-spouse	Machine-op-inspct	Husband
32558	Widowed	Adm-clerical	Unmarried
32559	Never-married	Adm-clerical	Own-child

```
32560      Married-civ-spouse      Exec-managerial           Wife
```

	race	sex	capital_gain	capital_loss	hpw \
0	White	Male	2174	0	40
1	White	Male	0	0	13
2	White	Male	0	0	40
3	Black	Male	0	0	40
4	Black	Female	0	0	40
5	White	Female	0	0	40
6	Black	Female	0	0	16
7	White	Male	0	0	45
8	White	Female	14084	0	50
9	White	Male	5178	0	40
10	Black	Male	0	0	80
11	Asian-Pac-Islander	Male	0	0	40
12	White	Female	0	0	30
13	Black	Male	0	0	50
14	Asian-Pac-Islander	Male	0	0	40
15	Amer-Indian-Eskimo	Male	0	0	45
16	White	Male	0	0	35
17	White	Male	0	0	40
18	White	Male	0	0	50
19	White	Female	0	0	45
20	White	Male	0	0	60
21	Black	Female	0	0	20
22	Black	Male	0	0	40
23	White	Male	0	2042	40
24	White	Female	0	0	40
25	White	Male	0	0	40
26	White	Male	0	0	40
27	Asian-Pac-Islander	Male	0	0	60
28	White	Male	0	0	80
29	White	Male	0	0	40
...	...	...	...	...	...
32531	Asian-Pac-Islander	Female	0	0	99
32532	White	Male	0	0	60
32533	Asian-Pac-Islander	Male	0	0	50
32534	White	Female	0	0	39
32535	Black	Male	0	0	35
32536	White	Female	0	0	55
32537	Black	Male	0	0	46
32538	Black	Female	15020	0	45
32539	White	Male	0	0	10
32540	White	Female	0	0	40
32541	Black	Female	0	0	32
32542	White	Male	0	0	25
32543	White	Female	0	0	48
32544	Other	Female	0	0	30

32545	White	Female	0	0	20
32546	White	Female	0	0	40
32547	White	Male	0	0	40
32548	White	Male	1086	0	60
32549	White	Female	0	0	40
32550	White	Male	0	0	50
32551	Amer-Indian-Eskimo	Male	0	0	40
32552	White	Male	0	0	45
32553	Asian-Pac-Islander	Male	0	0	11
32554	White	Male	0	0	40
32555	White	Male	0	0	40
32556	White	Female	0	0	38
32557	White	Male	0	0	40
32558	White	Female	0	0	40
32559	White	Male	0	0	20
32560	White	Female	15024	0	40

	native_country	salary_50K	workclass_num
0	United-States	<=50K	State-gov
1	United-States	<=50K	Self-emp-not-inc
2	United-States	<=50K	Private
3	United-States	<=50K	Private
4	Cuba	<=50K	Private
5	United-States	<=50K	Private
6	Jamaica	<=50K	Private
7	United-States	>50K	Self-emp-not-inc
8	United-States	>50K	Private
9	United-States	>50K	Private
10	United-States	>50K	Private
11	India	>50K	State-gov
12	United-States	<=50K	Private
13	United-States	<=50K	Private
14	?	>50K	Private
15	Mexico	<=50K	Private
16	United-States	<=50K	Self-emp-not-inc
17	United-States	<=50K	Private
18	United-States	<=50K	Private
19	United-States	>50K	Self-emp-not-inc
20	United-States	>50K	Private
21	United-States	<=50K	Private
22	United-States	<=50K	Federal-gov
23	United-States	<=50K	Private
24	United-States	<=50K	Private
25	United-States	>50K	Local-gov
26	United-States	<=50K	Private
27	South	>50K	?
28	United-States	<=50K	Private
29	United-States	<=50K	Private

```
            ...             ...   ...                 ...
32531    United-States   <=50K                   ?
32532    United-States    >50K             Private
32533            Japan    >50K             Private
32534    United-States   <=50K             Private
32535    United-States   <=50K             Private
32536    United-States    >50K             Private
32537    United-States   <=50K             Private
32538    United-States    >50K             Private
32539    United-States    >50K                   ?
32540    United-States   <=50K           State-gov
32541    United-States   <=50K                   ?
32542    United-States   <=50K                   ?
32543    United-States   <=50K           Local-gov
32544    United-States   <=50K             Private
32545    United-States    >50K           Local-gov
32546    United-States   <=50K             Private
32547           Mexico   <=50K             Private
32548    United-States   <=50K   Self-emp-not-inc
32549    United-States   <=50K           State-gov
32550    United-States   <=50K   Self-emp-not-inc
32551    United-States   <=50K             Private
32552    United-States   <=50K             Private
32553           Taiwan   <=50K             Private
32554    United-States    >50K             Private
32555    United-States   <=50K             Private
32556    United-States   <=50K             Private
32557    United-States    >50K             Private
32558    United-States   <=50K             Private
32559    United-States   <=50K             Private
32560    United-States    >50K        Self-emp-inc

[32561 rows x 16 columns]
```

In [24]: df['ms_num']= df['marital_status']
         df['occupation_num']= df['occupation']
         df['rel_num']= df['relationship']
         df['race_num']= df['race']
         df['sex_num']= df['sex']
         df['nc_num']= df['native_country']
         df

```
Out[24]:      age          workclass  fnlwgt   education  education_num  \
        0     39          State-gov   77516   Bachelors            13
        1     50   Self-emp-not-inc   83311   Bachelors            13
        2     38            Private  215646     HS-grad             9
        3     53            Private  234721        11th             7
        4     28            Private  338409   Bachelors            13
```

5	37	Private	284582	Masters	14
6	49	Private	160187	9th	5
7	52	Self-emp-not-inc	209642	HS-grad	9
8	31	Private	45781	Masters	14
9	42	Private	159449	Bachelors	13
10	37	Private	280464	Some-college	10
11	30	State-gov	141297	Bachelors	13
12	23	Private	122272	Bachelors	13
13	32	Private	205019	Assoc-acdm	12
14	40	Private	121772	Assoc-voc	11
15	34	Private	245487	7th-8th	4
16	25	Self-emp-not-inc	176756	HS-grad	9
17	32	Private	186824	HS-grad	9
18	38	Private	28887	11th	7
19	43	Self-emp-not-inc	292175	Masters	14
20	40	Private	193524	Doctorate	16
21	54	Private	302146	HS-grad	9
22	35	Federal-gov	76845	9th	5
23	43	Private	117037	11th	7
24	59	Private	109015	HS-grad	9
25	56	Local-gov	216851	Bachelors	13
26	19	Private	168294	HS-grad	9
27	54	?	180211	Some-college	10
28	39	Private	367260	HS-grad	9
29	49	Private	193366	HS-grad	9
...	...	...	...	...	...
32531	30	?	33811	Bachelors	13
32532	34	Private	204461	Doctorate	16
32533	54	Private	337992	Bachelors	13
32534	37	Private	179137	Some-college	10
32535	22	Private	325033	12th	8
32536	34	Private	160216	Bachelors	13
32537	30	Private	345898	HS-grad	9
32538	38	Private	139180	Bachelors	13
32539	71	?	287372	Doctorate	16
32540	45	State-gov	252208	HS-grad	9
32541	41	?	202822	HS-grad	9
32542	72	?	129912	HS-grad	9
32543	45	Local-gov	119199	Assoc-acdm	12
32544	31	Private	199655	Masters	14
32545	39	Local-gov	111499	Assoc-acdm	12
32546	37	Private	198216	Assoc-acdm	12
32547	43	Private	260761	HS-grad	9
32548	65	Self-emp-not-inc	99359	Prof-school	15
32549	43	State-gov	255835	Some-college	10
32550	43	Self-emp-not-inc	27242	Some-college	10
32551	32	Private	34066	10th	6
32552	43	Private	84661	Assoc-voc	11

```
32553  32        Private  116138       Masters            14
32554  53        Private  321865       Masters            14
32555  22        Private  310152  Some-college            10
32556  27        Private  257302    Assoc-acdm            12
32557  40        Private  154374       HS-grad             9
32558  58        Private  151910       HS-grad             9
32559  22        Private  201490       HS-grad             9
32560  52   Self-emp-inc  287927       HS-grad             9

              marital_status          occupation   relationship  \
0              Never-married        Adm-clerical  Not-in-family
1         Married-civ-spouse     Exec-managerial        Husband
2                   Divorced   Handlers-cleaners  Not-in-family
3         Married-civ-spouse   Handlers-cleaners        Husband
4         Married-civ-spouse      Prof-specialty           Wife
5         Married-civ-spouse     Exec-managerial           Wife
6       Married-spouse-absent      Other-service  Not-in-family
7         Married-civ-spouse     Exec-managerial        Husband
8              Never-married      Prof-specialty  Not-in-family
9         Married-civ-spouse     Exec-managerial        Husband
10        Married-civ-spouse     Exec-managerial        Husband
11        Married-civ-spouse      Prof-specialty        Husband
12             Never-married        Adm-clerical      Own-child
13             Never-married               Sales  Not-in-family
14        Married-civ-spouse         Craft-repair        Husband
15        Married-civ-spouse     Transport-moving        Husband
16             Never-married      Farming-fishing      Own-child
17             Never-married    Machine-op-inspct      Unmarried
18        Married-civ-spouse               Sales        Husband
19                  Divorced     Exec-managerial      Unmarried
20        Married-civ-spouse      Prof-specialty        Husband
21                 Separated       Other-service      Unmarried
22        Married-civ-spouse     Farming-fishing        Husband
23        Married-civ-spouse     Transport-moving        Husband
24                  Divorced        Tech-support      Unmarried
25        Married-civ-spouse        Tech-support        Husband
26             Never-married        Craft-repair      Own-child
27        Married-civ-spouse                   ?        Husband
28                  Divorced     Exec-managerial  Not-in-family
29        Married-civ-spouse        Craft-repair        Husband
...                      ...                 ...            ...
32531          Never-married                   ?  Not-in-family
32532     Married-civ-spouse      Prof-specialty        Husband
32533     Married-civ-spouse     Exec-managerial        Husband
32534               Divorced        Adm-clerical      Unmarried
32535          Never-married     Protective-serv      Own-child
32536          Never-married     Exec-managerial  Not-in-family
32537          Never-married        Craft-repair  Not-in-family
```

32538	Divorced	Prof-specialty	Unmarried	
32539	Married-civ-spouse	?	Husband	
32540	Separated	Adm-clerical	Own-child	
32541	Separated	?	Not-in-family	
32542	Married-civ-spouse	?	Husband	
32543	Divorced	Prof-specialty	Unmarried	
32544	Divorced	Other-service	Not-in-family	
32545	Married-civ-spouse	Adm-clerical	Wife	
32546	Divorced	Tech-support	Not-in-family	
32547	Married-civ-spouse	Machine-op-inspct	Husband	
32548	Never-married	Prof-specialty	Not-in-family	
32549	Divorced	Adm-clerical	Other-relative	
32550	Married-civ-spouse	Craft-repair	Husband	
32551	Married-civ-spouse	Handlers-cleaners	Husband	
32552	Married-civ-spouse	Sales	Husband	
32553	Never-married	Tech-support	Not-in-family	
32554	Married-civ-spouse	Exec-managerial	Husband	
32555	Never-married	Protective-serv	Not-in-family	
32556	Married-civ-spouse	Tech-support	Wife	
32557	Married-civ-spouse	Machine-op-inspct	Husband	
32558	Widowed	Adm-clerical	Unmarried	
32559	Never-married	Adm-clerical	Own-child	
32560	Married-civ-spouse	Exec-managerial	Wife	

	race	sex	...	hpw	native_country \
0	White	Male	...	40	United-States
1	White	Male	...	13	United-States
2	White	Male	...	40	United-States
3	Black	Male	...	40	United-States
4	Black	Female	...	40	Cuba
5	White	Female	...	40	United-States
6	Black	Female	...	16	Jamaica
7	White	Male	...	45	United-States
8	White	Female	...	50	United-States
9	White	Male	...	40	United-States
10	Black	Male	...	80	United-States
11	Asian-Pac-Islander	Male	...	40	India
12	White	Female	...	30	United-States
13	Black	Male	...	50	United-States
14	Asian-Pac-Islander	Male	...	40	?
15	Amer-Indian-Eskimo	Male	...	45	Mexico
16	White	Male	...	35	United-States
17	White	Male	...	40	United-States
18	White	Male	...	50	United-States
19	White	Female	...	45	United-States
20	White	Male	...	60	United-States
21	Black	Female	...	20	United-States
22	Black	Male	...	40	United-States

```
23                    White      Male    ...        40     United-States
24                    White    Female    ...        40     United-States
25                    White      Male    ...        40     United-States
26                    White      Male    ...        40     United-States
27      Asian-Pac-Islander      Male    ...        60             South
28                    White      Male    ...        80     United-States
29                    White      Male    ...        40     United-States
...                     ...       ...    ...       ...               ...
32531   Asian-Pac-Islander    Female    ...        99     United-States
32532                 White      Male    ...        60     United-States
32533   Asian-Pac-Islander      Male    ...        50             Japan
32534                 White    Female    ...        39     United-States
32535                 Black      Male    ...        35     United-States
32536                 White    Female    ...        55     United-States
32537                 Black      Male    ...        46     United-States
32538                 Black    Female    ...        45     United-States
32539                 White      Male    ...        10     United-States
32540                 White    Female    ...        40     United-States
32541                 Black    Female    ...        32     United-States
32542                 White      Male    ...        25     United-States
32543                 White    Female    ...        48     United-States
32544                 Other    Female    ...        30     United-States
32545                 White    Female    ...        20     United-States
32546                 White    Female    ...        40     United-States
32547                 White      Male    ...        40            Mexico
32548                 White      Male    ...        60     United-States
32549                 White    Female    ...        40     United-States
32550                 White      Male    ...        50     United-States
32551    Amer-Indian-Eskimo      Male    ...        40     United-States
32552                 White      Male    ...        45     United-States
32553   Asian-Pac-Islander      Male    ...        11            Taiwan
32554                 White      Male    ...        40     United-States
32555                 White      Male    ...        40     United-States
32556                 White    Female    ...        38     United-States
32557                 White      Male    ...        40     United-States
32558                 White    Female    ...        40     United-States
32559                 White      Male    ...        20     United-States
32560                 White    Female    ...        40     United-States


       salary_50K      workclass_num                 ms_num  \
0           <=50K          State-gov          Never-married
1           <=50K   Self-emp-not-inc     Married-civ-spouse
2           <=50K            Private               Divorced
3           <=50K            Private     Married-civ-spouse
4           <=50K            Private     Married-civ-spouse
5           <=50K            Private     Married-civ-spouse
6           <=50K            Private  Married-spouse-absent
7            >50K   Self-emp-not-inc     Married-civ-spouse
```

8	>50K	Private	Never-married
9	>50K	Private	Married-civ-spouse
10	>50K	Private	Married-civ-spouse
11	>50K	State-gov	Married-civ-spouse
12	<=50K	Private	Never-married
13	<=50K	Private	Never-married
14	>50K	Private	Married-civ-spouse
15	<=50K	Private	Married-civ-spouse
16	<=50K	Self-emp-not-inc	Never-married
17	<=50K	Private	Never-married
18	<=50K	Private	Married-civ-spouse
19	>50K	Self-emp-not-inc	Divorced
20	>50K	Private	Married-civ-spouse
21	<=50K	Private	Separated
22	<=50K	Federal-gov	Married-civ-spouse
23	<=50K	Private	Married-civ-spouse
24	<=50K	Private	Divorced
25	>50K	Local-gov	Married-civ-spouse
26	<=50K	Private	Never-married
27	>50K	?	Married-civ-spouse
28	<=50K	Private	Divorced
29	<=50K	Private	Married-civ-spouse
...	...	...	...
32531	<=50K	?	Never-married
32532	>50K	Private	Married-civ-spouse
32533	>50K	Private	Married-civ-spouse
32534	<=50K	Private	Divorced
32535	<=50K	Private	Never-married
32536	>50K	Private	Never-married
32537	<=50K	Private	Never-married
32538	>50K	Private	Divorced
32539	>50K	?	Married-civ-spouse
32540	<=50K	State-gov	Separated
32541	<=50K	?	Separated
32542	<=50K	?	Married-civ-spouse
32543	<=50K	Local-gov	Divorced
32544	<=50K	Private	Divorced
32545	>50K	Local-gov	Married-civ-spouse
32546	<=50K	Private	Divorced
32547	<=50K	Private	Married-civ-spouse
32548	<=50K	Self-emp-not-inc	Never-married
32549	<=50K	State-gov	Divorced
32550	<=50K	Self-emp-not-inc	Married-civ-spouse
32551	<=50K	Private	Married-civ-spouse
32552	<=50K	Private	Married-civ-spouse
32553	<=50K	Private	Never-married
32554	>50K	Private	Married-civ-spouse
32555	<=50K	Private	Never-married

```
32556     <=50K          Private     Married-civ-spouse
32557      >50K          Private     Married-civ-spouse
32558     <=50K          Private                Widowed
32559     <=50K          Private          Never-married
32560      >50K      Self-emp-inc     Married-civ-spouse


          occupation_num        rel_num            race_num  sex_num  \
0            Adm-clerical    Not-in-family            White     Male
1         Exec-managerial          Husband            White     Male
2       Handlers-cleaners    Not-in-family            White     Male
3       Handlers-cleaners          Husband            Black     Male
4           Prof-specialty             Wife            Black   Female
5         Exec-managerial             Wife            White   Female
6           Other-service    Not-in-family            Black   Female
7         Exec-managerial          Husband            White     Male
8           Prof-specialty    Not-in-family            White   Female
9         Exec-managerial          Husband            White     Male
10        Exec-managerial          Husband            Black     Male
11          Prof-specialty          Husband  Asian-Pac-Islander   Male
12           Adm-clerical        Own-child            White   Female
13                 Sales    Not-in-family            Black     Male
14           Craft-repair          Husband  Asian-Pac-Islander   Male
15       Transport-moving          Husband  Amer-Indian-Eskimo   Male
16         Farming-fishing        Own-child            White     Male
17      Machine-op-inspct        Unmarried            White     Male
18                 Sales          Husband            White     Male
19        Exec-managerial        Unmarried            White   Female
20          Prof-specialty          Husband            White     Male
21          Other-service        Unmarried            Black   Female
22         Farming-fishing          Husband            Black     Male
23       Transport-moving          Husband            White     Male
24           Tech-support        Unmarried            White   Female
25           Tech-support          Husband            White     Male
26           Craft-repair        Own-child            White     Male
27                     ?          Husband  Asian-Pac-Islander   Male
28        Exec-managerial    Not-in-family            White     Male
29           Craft-repair          Husband            White     Male
...                  ...              ...                ...      ...
32531                  ?    Not-in-family  Asian-Pac-Islander  Female
32532       Prof-specialty          Husband            White     Male
32533     Exec-managerial          Husband  Asian-Pac-Islander   Male
32534        Adm-clerical        Unmarried            White   Female
32535     Protective-serv        Own-child            Black     Male
32536     Exec-managerial    Not-in-family            White   Female
32537        Craft-repair    Not-in-family            Black     Male
32538       Prof-specialty        Unmarried            Black   Female
32539                  ?          Husband            White     Male
32540        Adm-clerical        Own-child            White   Female
```

```
32541                 ?  Not-in-family            Black  Female
32542                 ?      Husband            White    Male
32543     Prof-specialty    Unmarried            White  Female
32544      Other-service  Not-in-family            Other  Female
32545       Adm-clerical         Wife            White  Female
32546       Tech-support  Not-in-family            White  Female
32547  Machine-op-inspct      Husband            White    Male
32548     Prof-specialty  Not-in-family            White    Male
32549       Adm-clerical  Other-relative            White  Female
32550       Craft-repair      Husband            White    Male
32551  Handlers-cleaners      Husband  Amer-Indian-Eskimo    Male
32552              Sales      Husband            White    Male
32553       Tech-support  Not-in-family  Asian-Pac-Islander    Male
32554    Exec-managerial      Husband            White    Male
32555    Protective-serv  Not-in-family            White    Male
32556       Tech-support         Wife            White  Female
32557  Machine-op-inspct      Husband            White    Male
32558       Adm-clerical    Unmarried            White  Female
32559       Adm-clerical    Own-child            White    Male
32560    Exec-managerial         Wife            White  Female

            nc_num
0      United-States
1      United-States
2      United-States
3      United-States
4               Cuba
5      United-States
6            Jamaica
7      United-States
8      United-States
9      United-States
10     United-States
11             India
12     United-States
13     United-States
14                 ?
15            Mexico
16     United-States
17     United-States
18     United-States
19     United-States
20     United-States
21     United-States
22     United-States
23     United-States
24     United-States
25     United-States
```

```
26      United-States
27              South
28      United-States
29      United-States
...               ...
32531   United-States
32532   United-States
32533           Japan
32534   United-States
32535   United-States
32536   United-States
32537   United-States
32538   United-States
32539   United-States
32540   United-States
32541   United-States
32542   United-States
32543   United-States
32544   United-States
32545   United-States
32546   United-States
32547          Mexico
32548   United-States
32549   United-States
32550   United-States
32551   United-States
32552   United-States
32553          Taiwan
32554   United-States
32555   United-States
32556   United-States
32557   United-States
32558   United-States
32559   United-States
32560   United-States

[32561 rows x 22 columns]

In [26]: df_with_dummies = pd.get_dummies(df, prefix='workclass_num', columns=['workclass_num']

In [27]: df_with_dummies

Out[27]:       age         workclass  fnlwgt    education  education_num  \
         0      39          State-gov   77516    Bachelors             13
         1      50   Self-emp-not-inc   83311    Bachelors             13
         2      38            Private  215646      HS-grad              9
         3      53            Private  234721         11th              7
         4      28            Private  338409    Bachelors             13
```

5	37	Private	284582	Masters	14
6	49	Private	160187	9th	5
7	52	Self-emp-not-inc	209642	HS-grad	9
8	31	Private	45781	Masters	14
9	42	Private	159449	Bachelors	13
10	37	Private	280464	Some-college	10
11	30	State-gov	141297	Bachelors	13
12	23	Private	122272	Bachelors	13
13	32	Private	205019	Assoc-acdm	12
14	40	Private	121772	Assoc-voc	11
15	34	Private	245487	7th-8th	4
16	25	Self-emp-not-inc	176756	HS-grad	9
17	32	Private	186824	HS-grad	9
18	38	Private	28887	11th	7
19	43	Self-emp-not-inc	292175	Masters	14
20	40	Private	193524	Doctorate	16
21	54	Private	302146	HS-grad	9
22	35	Federal-gov	76845	9th	5
23	43	Private	117037	11th	7
24	59	Private	109015	HS-grad	9
25	56	Local-gov	216851	Bachelors	13
26	19	Private	168294	HS-grad	9
27	54	?	180211	Some-college	10
28	39	Private	367260	HS-grad	9
29	49	Private	193366	HS-grad	9
...	...	...	...	...	...
32531	30	?	33811	Bachelors	13
32532	34	Private	204461	Doctorate	16
32533	54	Private	337992	Bachelors	13
32534	37	Private	179137	Some-college	10
32535	22	Private	325033	12th	8
32536	34	Private	160216	Bachelors	13
32537	30	Private	345898	HS-grad	9
32538	38	Private	139180	Bachelors	13
32539	71	?	287372	Doctorate	16
32540	45	State-gov	252208	HS-grad	9
32541	41	?	202822	HS-grad	9
32542	72	?	129912	HS-grad	9
32543	45	Local-gov	119199	Assoc-acdm	12
32544	31	Private	199655	Masters	14
32545	39	Local-gov	111499	Assoc-acdm	12
32546	37	Private	198216	Assoc-acdm	12
32547	43	Private	260761	HS-grad	9
32548	65	Self-emp-not-inc	99359	Prof-school	15
32549	43	State-gov	255835	Some-college	10
32550	43	Self-emp-not-inc	27242	Some-college	10
32551	32	Private	34066	10th	6
32552	43	Private	84661	Assoc-voc	11

```
32553  32        Private  116138      Masters         14
32554  53        Private  321865      Masters         14
32555  22        Private  310152  Some-college        10
32556  27        Private  257302    Assoc-acdm        12
32557  40        Private  154374       HS-grad          9
32558  58        Private  151910       HS-grad          9
32559  22        Private  201490       HS-grad          9
32560  52   Self-emp-inc  287927       HS-grad          9

                 marital_status           occupation   relationship  \
0                 Never-married         Adm-clerical   Not-in-family
1            Married-civ-spouse      Exec-managerial        Husband
2                      Divorced    Handlers-cleaners   Not-in-family
3            Married-civ-spouse    Handlers-cleaners        Husband
4            Married-civ-spouse       Prof-specialty           Wife
5            Married-civ-spouse      Exec-managerial           Wife
6          Married-spouse-absent        Other-service   Not-in-family
7            Married-civ-spouse      Exec-managerial        Husband
8                 Never-married       Prof-specialty   Not-in-family
9            Married-civ-spouse      Exec-managerial        Husband
10           Married-civ-spouse      Exec-managerial        Husband
11           Married-civ-spouse       Prof-specialty        Husband
12                Never-married         Adm-clerical      Own-child
13                Never-married                Sales   Not-in-family
14           Married-civ-spouse          Craft-repair        Husband
15           Married-civ-spouse      Transport-moving        Husband
16                Never-married       Farming-fishing      Own-child
17                Never-married     Machine-op-inspct      Unmarried
18           Married-civ-spouse                Sales        Husband
19                     Divorced      Exec-managerial      Unmarried
20           Married-civ-spouse       Prof-specialty        Husband
21                    Separated        Other-service      Unmarried
22           Married-civ-spouse       Farming-fishing        Husband
23           Married-civ-spouse      Transport-moving        Husband
24                     Divorced         Tech-support      Unmarried
25           Married-civ-spouse         Tech-support        Husband
26                Never-married          Craft-repair      Own-child
27           Married-civ-spouse                    ?        Husband
28                     Divorced      Exec-managerial   Not-in-family
29           Married-civ-spouse          Craft-repair        Husband
...                         ...                  ...            ...
32531             Never-married                    ?   Not-in-family
32532        Married-civ-spouse       Prof-specialty        Husband
32533        Married-civ-spouse      Exec-managerial        Husband
32534                  Divorced         Adm-clerical      Unmarried
32535             Never-married      Protective-serv      Own-child
32536             Never-married      Exec-managerial   Not-in-family
32537             Never-married          Craft-repair   Not-in-family
```

```
32538                Divorced         Prof-specialty       Unmarried
32539      Married-civ-spouse                      ?         Husband
32540               Separated          Adm-clerical       Own-child
32541               Separated                      ?    Not-in-family
32542      Married-civ-spouse                      ?         Husband
32543                Divorced         Prof-specialty       Unmarried
32544                Divorced          Other-service    Not-in-family
32545      Married-civ-spouse          Adm-clerical            Wife
32546                Divorced           Tech-support    Not-in-family
32547      Married-civ-spouse      Machine-op-inspct         Husband
32548           Never-married         Prof-specialty    Not-in-family
32549                Divorced          Adm-clerical    Other-relative
32550      Married-civ-spouse           Craft-repair         Husband
32551      Married-civ-spouse      Handlers-cleaners         Husband
32552      Married-civ-spouse                  Sales         Husband
32553           Never-married           Tech-support    Not-in-family
32554      Married-civ-spouse        Exec-managerial         Husband
32555           Never-married        Protective-serv    Not-in-family
32556      Married-civ-spouse           Tech-support            Wife
32557      Married-civ-spouse      Machine-op-inspct         Husband
32558                 Widowed          Adm-clerical       Unmarried
32559           Never-married          Adm-clerical       Own-child
32560      Married-civ-spouse        Exec-managerial            Wife

                          race      sex          ...                    \
0                        White     Male          ...
1                        White     Male          ...
2                        White     Male          ...
3                        Black     Male          ...
4                        Black   Female          ...
5                        White   Female          ...
6                        Black   Female          ...
7                        White     Male          ...
8                        White   Female          ...
9                        White     Male          ...
10                       Black     Male          ...
11          Asian-Pac-Islander     Male          ...
12                       White   Female          ...
13                       Black     Male          ...
14          Asian-Pac-Islander     Male          ...
15          Amer-Indian-Eskimo     Male          ...
16                       White     Male          ...
17                       White     Male          ...
18                       White     Male          ...
19                       White   Female          ...
20                       White     Male          ...
21                       Black   Female          ...
22                       Black     Male          ...
```

```
23                 White    Male      ...
24                 White  Female      ...
25                 White    Male      ...
26                 White    Male      ...
27     Asian-Pac-Islander    Male      ...
28                 White    Male      ...
29                 White    Male      ...
...                  ...     ...      ...
32531  Asian-Pac-Islander  Female      ...
32532              White    Male      ...
32533  Asian-Pac-Islander    Male      ...
32534              White  Female      ...
32535              Black    Male      ...
32536              White  Female      ...
32537              Black    Male      ...
32538              Black  Female      ...
32539              White    Male      ...
32540              White  Female      ...
32541              Black  Female      ...
32542              White    Male      ...
32543              White  Female      ...
32544              Other  Female      ...
32545              White  Female      ...
32546              White  Female      ...
32547              White    Male      ...
32548              White    Male      ...
32549              White  Female      ...
32550              White    Male      ...
32551   Amer-Indian-Eskimo    Male      ...
32552              White    Male      ...
32553  Asian-Pac-Islander    Male      ...
32554              White    Male      ...
32555              White    Male      ...
32556              White  Female      ...
32557              White    Male      ...
32558              White  Female      ...
32559              White    Male      ...
32560              White  Female      ...


             nc_num  workclass_num_ ?  workclass_num_ Federal-gov  \
0      United-States                0                           0
1      United-States                0                           0
2      United-States                0                           0
3      United-States                0                           0
4               Cuba                0                           0
5      United-States                0                           0
6            Jamaica                0                           0
7      United-States                0                           0
```

8	United-States	0	0
9	United-States	0	0
10	United-States	0	0
11	India	0	0
12	United-States	0	0
13	United-States	0	0
14	?	0	0
15	Mexico	0	0
16	United-States	0	0
17	United-States	0	0
18	United-States	0	0
19	United-States	0	0
20	United-States	0	0
21	United-States	0	0
22	United-States	0	1
23	United-States	0	0
24	United-States	0	0
25	United-States	0	0
26	United-States	0	0
27	South	1	0
28	United-States	0	0
29	United-States	0	0
...	...	...	...
32531	United-States	1	0
32532	United-States	0	0
32533	Japan	0	0
32534	United-States	0	0
32535	United-States	0	0
32536	United-States	0	0
32537	United-States	0	0
32538	United-States	0	0
32539	United-States	1	0
32540	United-States	0	0
32541	United-States	1	0
32542	United-States	1	0
32543	United-States	0	0
32544	United-States	0	0
32545	United-States	0	0
32546	United-States	0	0
32547	Mexico	0	0
32548	United-States	0	0
32549	United-States	0	0
32550	United-States	0	0
32551	United-States	0	0
32552	United-States	0	0
32553	Taiwan	0	0
32554	United-States	0	0
32555	United-States	0	0

```
32556    United-States              0                          0
32557    United-States              0                          0
32558    United-States              0                          0
32559    United-States              0                          0
32560    United-States              0                          0

       workclass_num_ Local-gov workclass_num_ Never-worked  \
0                            0                          0
1                            0                          0
2                            0                          0
3                            0                          0
4                            0                          0
5                            0                          0
6                            0                          0
7                            0                          0
8                            0                          0
9                            0                          0
10                           0                          0
11                           0                          0
12                           0                          0
13                           0                          0
14                           0                          0
15                           0                          0
16                           0                          0
17                           0                          0
18                           0                          0
19                           0                          0
20                           0                          0
21                           0                          0
22                           0                          0
23                           0                          0
24                           0                          0
25                           1                          0
26                           0                          0
27                           0                          0
28                           0                          0
29                           0                          0
...                        ...                        ...
32531                        0                          0
32532                        0                          0
32533                        0                          0
32534                        0                          0
32535                        0                          0
32536                        0                          0
32537                        0                          0
32538                        0                          0
32539                        0                          0
32540                        0                          0
```

32541	0	0
32542	0	0
32543	1	0
32544	0	0
32545	1	0
32546	0	0
32547	0	0
32548	0	0
32549	0	0
32550	0	0
32551	0	0
32552	0	0
32553	0	0
32554	0	0
32555	0	0
32556	0	0
32557	0	0
32558	0	0
32559	0	0
32560	0	0

	workclass_num_ Private	workclass_num_ Self-emp-inc \
0	0	0
1	0	0
2	1	0
3	1	0
4	1	0
5	1	0
6	1	0
7	0	0
8	1	0
9	1	0
10	1	0
11	0	0
12	1	0
13	1	0
14	1	0
15	1	0
16	0	0
17	1	0
18	1	0
19	0	0
20	1	0
21	1	0
22	0	0
23	1	0
24	1	0
25	0	0

26	1	0
27	0	0
28	1	0
29	1	0
...	...	...
32531	0	0
32532	1	0
32533	1	0
32534	1	0
32535	1	0
32536	1	0
32537	1	0
32538	1	0
32539	0	0
32540	0	0
32541	0	0
32542	0	0
32543	0	0
32544	1	0
32545	0	0
32546	1	0
32547	1	0
32548	0	0
32549	0	0
32550	0	0
32551	1	0
32552	1	0
32553	1	0
32554	1	0
32555	1	0
32556	1	0
32557	1	0
32558	1	0
32559	1	0
32560	0	1

	workclass_num_ Self-emp-not-inc	workclass_num_ State-gov \
0	0	1
1	1	0
2	0	0
3	0	0
4	0	0
5	0	0
6	0	0
7	1	0
8	0	0
9	0	0
10	0	0

11	0	1
12	0	0
13	0	0
14	0	0
15	0	0
16	1	0
17	0	0
18	0	0
19	1	0
20	0	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	0	0
29	0	0
...	...	...
32531	0	0
32532	0	0
32533	0	0
32534	0	0
32535	0	0
32536	0	0
32537	0	0
32538	0	0
32539	0	0
32540	0	1
32541	0	0
32542	0	0
32543	0	0
32544	0	0
32545	0	0
32546	0	0
32547	0	0
32548	1	0
32549	0	1
32550	1	0
32551	0	0
32552	0	0
32553	0	0
32554	0	0
32555	0	0
32556	0	0
32557	0	0
32558	0	0

32559	0	0
32560	0	0

	workclass_num_ Without-pay
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0
17	0
18	0
19	0
20	0
21	0
22	0
23	0
24	0
25	0
26	0
27	0
28	0
29	0
...	...
32531	0
32532	0
32533	0
32534	0
32535	0
32536	0
32537	0
32538	0
32539	0
32540	0
32541	0
32542	0
32543	0

```
32544                          0
32545                          0
32546                          0
32547                          0
32548                          0
32549                          0
32550                          0
32551                          0
32552                          0
32553                          0
32554                          0
32555                          0
32556                          0
32557                          0
32558                          0
32559                          0
32560                          0

[32561 rows x 30 columns]

In [28]: df

Out[28]:        age          workclass  fnlwgt     education  education_num  \
        0        39           State-gov   77516      Bachelors             13
        1        50    Self-emp-not-inc   83311      Bachelors             13
        2        38             Private  215646        HS-grad              9
        3        53             Private  234721           11th              7
        4        28             Private  338409      Bachelors             13
        5        37             Private  284582        Masters             14
        6        49             Private  160187            9th              5
        7        52    Self-emp-not-inc  209642        HS-grad              9
        8        31             Private   45781        Masters             14
        9        42             Private  159449      Bachelors             13
        10       37             Private  280464   Some-college             10
        11       30           State-gov  141297      Bachelors             13
        12       23             Private  122272      Bachelors             13
        13       32             Private  205019     Assoc-acdm             12
        14       40             Private  121772      Assoc-voc             11
        15       34             Private  245487        7th-8th              4
        16       25    Self-emp-not-inc  176756        HS-grad              9
        17       32             Private  186824        HS-grad              9
        18       38             Private   28887           11th              7
        19       43    Self-emp-not-inc  292175        Masters             14
        20       40             Private  193524      Doctorate             16
        21       54             Private  302146        HS-grad              9
        22       35         Federal-gov   76845            9th              5
        23       43             Private  117037           11th              7
        24       59             Private  109015        HS-grad              9
```

25	56	Local-gov	216851	Bachelors	13
26	19	Private	168294	HS-grad	9
27	54	?	180211	Some-college	10
28	39	Private	367260	HS-grad	9
29	49	Private	193366	HS-grad	9
...	...	...	...	...	...
32531	30	?	33811	Bachelors	13
32532	34	Private	204461	Doctorate	16
32533	54	Private	337992	Bachelors	13
32534	37	Private	179137	Some-college	10
32535	22	Private	325033	12th	8
32536	34	Private	160216	Bachelors	13
32537	30	Private	345898	HS-grad	9
32538	38	Private	139180	Bachelors	13
32539	71	?	287372	Doctorate	16
32540	45	State-gov	252208	HS-grad	9
32541	41	?	202822	HS-grad	9
32542	72	?	129912	HS-grad	9
32543	45	Local-gov	119199	Assoc-acdm	12
32544	31	Private	199655	Masters	14
32545	39	Local-gov	111499	Assoc-acdm	12
32546	37	Private	198216	Assoc-acdm	12
32547	43	Private	260761	HS-grad	9
32548	65	Self-emp-not-inc	99359	Prof-school	15
32549	43	State-gov	255835	Some-college	10
32550	43	Self-emp-not-inc	27242	Some-college	10
32551	32	Private	34066	10th	6
32552	43	Private	84661	Assoc-voc	11
32553	32	Private	116138	Masters	14
32554	53	Private	321865	Masters	14
32555	22	Private	310152	Some-college	10
32556	27	Private	257302	Assoc-acdm	12
32557	40	Private	154374	HS-grad	9
32558	58	Private	151910	HS-grad	9
32559	22	Private	201490	HS-grad	9
32560	52	Self-emp-inc	287927	HS-grad	9

	marital_status	occupation	relationship \
0	Never-married	Adm-clerical	Not-in-family
1	Married-civ-spouse	Exec-managerial	Husband
2	Divorced	Handlers-cleaners	Not-in-family
3	Married-civ-spouse	Handlers-cleaners	Husband
4	Married-civ-spouse	Prof-specialty	Wife
5	Married-civ-spouse	Exec-managerial	Wife
6	Married-spouse-absent	Other-service	Not-in-family
7	Married-civ-spouse	Exec-managerial	Husband
8	Never-married	Prof-specialty	Not-in-family
9	Married-civ-spouse	Exec-managerial	Husband

10	Married-civ-spouse	Exec-managerial	Husband
11	Married-civ-spouse	Prof-specialty	Husband
12	Never-married	Adm-clerical	Own-child
13	Never-married	Sales	Not-in-family
14	Married-civ-spouse	Craft-repair	Husband
15	Married-civ-spouse	Transport-moving	Husband
16	Never-married	Farming-fishing	Own-child
17	Never-married	Machine-op-inspct	Unmarried
18	Married-civ-spouse	Sales	Husband
19	Divorced	Exec-managerial	Unmarried
20	Married-civ-spouse	Prof-specialty	Husband
21	Separated	Other-service	Unmarried
22	Married-civ-spouse	Farming-fishing	Husband
23	Married-civ-spouse	Transport-moving	Husband
24	Divorced	Tech-support	Unmarried
25	Married-civ-spouse	Tech-support	Husband
26	Never-married	Craft-repair	Own-child
27	Married-civ-spouse	?	Husband
28	Divorced	Exec-managerial	Not-in-family
29	Married-civ-spouse	Craft-repair	Husband
...	...	...	...
32531	Never-married	?	Not-in-family
32532	Married-civ-spouse	Prof-specialty	Husband
32533	Married-civ-spouse	Exec-managerial	Husband
32534	Divorced	Adm-clerical	Unmarried
32535	Never-married	Protective-serv	Own-child
32536	Never-married	Exec-managerial	Not-in-family
32537	Never-married	Craft-repair	Not-in-family
32538	Divorced	Prof-specialty	Unmarried
32539	Married-civ-spouse	?	Husband
32540	Separated	Adm-clerical	Own-child
32541	Separated	?	Not-in-family
32542	Married-civ-spouse	?	Husband
32543	Divorced	Prof-specialty	Unmarried
32544	Divorced	Other-service	Not-in-family
32545	Married-civ-spouse	Adm-clerical	Wife
32546	Divorced	Tech-support	Not-in-family
32547	Married-civ-spouse	Machine-op-inspct	Husband
32548	Never-married	Prof-specialty	Not-in-family
32549	Divorced	Adm-clerical	Other-relative
32550	Married-civ-spouse	Craft-repair	Husband
32551	Married-civ-spouse	Handlers-cleaners	Husband
32552	Married-civ-spouse	Sales	Husband
32553	Never-married	Tech-support	Not-in-family
32554	Married-civ-spouse	Exec-managerial	Husband
32555	Never-married	Protective-serv	Not-in-family
32556	Married-civ-spouse	Tech-support	Wife
32557	Married-civ-spouse	Machine-op-inspct	Husband

			...		
32558	Widowed	Adm-clerical		Unmarried	
32559	Never-married	Adm-clerical		Own-child	
32560	Married-civ-spouse	Exec-managerial		Wife	

	race	sex	...	hpw	native_country \
0	White	Male	...	40	United-States
1	White	Male	...	13	United-States
2	White	Male	...	40	United-States
3	Black	Male	...	40	United-States
4	Black	Female	...	40	Cuba
5	White	Female	...	40	United-States
6	Black	Female	...	16	Jamaica
7	White	Male	...	45	United-States
8	White	Female	...	50	United-States
9	White	Male	...	40	United-States
10	Black	Male	...	80	United-States
11	Asian-Pac-Islander	Male	...	40	India
12	White	Female	...	30	United-States
13	Black	Male	...	50	United-States
14	Asian-Pac-Islander	Male	...	40	?
15	Amer-Indian-Eskimo	Male	...	45	Mexico
16	White	Male	...	35	United-States
17	White	Male	...	40	United-States
18	White	Male	...	50	United-States
19	White	Female	...	45	United-States
20	White	Male	...	60	United-States
21	Black	Female	...	20	United-States
22	Black	Male	...	40	United-States
23	White	Male	...	40	United-States
24	White	Female	...	40	United-States
25	White	Male	...	40	United-States
26	White	Male	...	40	United-States
27	Asian-Pac-Islander	Male	...	60	South
28	White	Male	...	80	United-States
29	White	Male	...	40	United-States
...	...	...	...	...	...
32531	Asian-Pac-Islander	Female	...	99	United-States
32532	White	Male	...	60	United-States
32533	Asian-Pac-Islander	Male	...	50	Japan
32534	White	Female	...	39	United-States
32535	Black	Male	...	35	United-States
32536	White	Female	...	55	United-States
32537	Black	Male	...	46	United-States
32538	Black	Female	...	45	United-States
32539	White	Male	...	10	United-States
32540	White	Female	...	40	United-States
32541	Black	Female	...	32	United-States
32542	White	Male	...	25	United-States

```
32543                       White    Female        ...          48     United-States
32544                       Other    Female        ...          30     United-States
32545                       White    Female        ...          20     United-States
32546                       White    Female        ...          40     United-States
32547                       White      Male        ...          40            Mexico
32548                       White      Male        ...          60     United-States
32549                       White    Female        ...          40     United-States
32550                       White      Male        ...          50     United-States
32551   Amer-Indian-Eskimo    Male        ...          40     United-States
32552                       White      Male        ...          45     United-States
32553   Asian-Pac-Islander    Male        ...          11            Taiwan
32554                       White      Male        ...          40     United-States
32555                       White      Male        ...          40     United-States
32556                       White    Female        ...          38     United-States
32557                       White      Male        ...          40     United-States
32558                       White    Female        ...          40     United-States
32559                       White      Male        ...          20     United-States
32560                       White    Female        ...          40     United-States


        salary_50K        workclass_num                    ms_num  \
0           <=50K              State-gov         Never-married
1           <=50K      Self-emp-not-inc    Married-civ-spouse
2           <=50K                Private              Divorced
3           <=50K                Private    Married-civ-spouse
4           <=50K                Private    Married-civ-spouse
5           <=50K                Private    Married-civ-spouse
6           <=50K                Private  Married-spouse-absent
7            >50K      Self-emp-not-inc    Married-civ-spouse
8            >50K                Private         Never-married
9            >50K                Private    Married-civ-spouse
10           >50K                Private    Married-civ-spouse
11           >50K              State-gov    Married-civ-spouse
12          <=50K                Private         Never-married
13          <=50K                Private         Never-married
14           >50K                Private    Married-civ-spouse
15          <=50K                Private    Married-civ-spouse
16          <=50K      Self-emp-not-inc         Never-married
17          <=50K                Private         Never-married
18          <=50K                Private    Married-civ-spouse
19           >50K      Self-emp-not-inc              Divorced
20           >50K                Private    Married-civ-spouse
21          <=50K                Private             Separated
22          <=50K            Federal-gov    Married-civ-spouse
23          <=50K                Private    Married-civ-spouse
24          <=50K                Private              Divorced
25           >50K              Local-gov    Married-civ-spouse
26          <=50K                Private         Never-married
27           >50K                    ?    Married-civ-spouse
```

28	<=50K	Private	Divorced
29	<=50K	Private	Married-civ-spouse
...	...	...	...
32531	<=50K	?	Never-married
32532	>50K	Private	Married-civ-spouse
32533	>50K	Private	Married-civ-spouse
32534	<=50K	Private	Divorced
32535	<=50K	Private	Never-married
32536	>50K	Private	Never-married
32537	<=50K	Private	Never-married
32538	>50K	Private	Divorced
32539	>50K	?	Married-civ-spouse
32540	<=50K	State-gov	Separated
32541	<=50K	?	Separated
32542	<=50K	?	Married-civ-spouse
32543	<=50K	Local-gov	Divorced
32544	<=50K	Private	Divorced
32545	>50K	Local-gov	Married-civ-spouse
32546	<=50K	Private	Divorced
32547	<=50K	Private	Married-civ-spouse
32548	<=50K	Self-emp-not-inc	Never-married
32549	<=50K	State-gov	Divorced
32550	<=50K	Self-emp-not-inc	Married-civ-spouse
32551	<=50K	Private	Married-civ-spouse
32552	<=50K	Private	Married-civ-spouse
32553	<=50K	Private	Never-married
32554	>50K	Private	Married-civ-spouse
32555	<=50K	Private	Never-married
32556	<=50K	Private	Married-civ-spouse
32557	>50K	Private	Married-civ-spouse
32558	<=50K	Private	Widowed
32559	<=50K	Private	Never-married
32560	>50K	Self-emp-inc	Married-civ-spouse

	occupation_num	rel_num	race_num	sex_num	\
0	Adm-clerical	Not-in-family	White	Male	
1	Exec-managerial	Husband	White	Male	
2	Handlers-cleaners	Not-in-family	White	Male	
3	Handlers-cleaners	Husband	Black	Male	
4	Prof-specialty	Wife	Black	Female	
5	Exec-managerial	Wife	White	Female	
6	Other-service	Not-in-family	Black	Female	
7	Exec-managerial	Husband	White	Male	
8	Prof-specialty	Not-in-family	White	Female	
9	Exec-managerial	Husband	White	Male	
10	Exec-managerial	Husband	Black	Male	
11	Prof-specialty	Husband	Asian-Pac-Islander	Male	
12	Adm-clerical	Own-child	White	Female	

13	Sales	Not-in-family	Black	Male
14	Craft-repair	Husband	Asian-Pac-Islander	Male
15	Transport-moving	Husband	Amer-Indian-Eskimo	Male
16	Farming-fishing	Own-child	White	Male
17	Machine-op-inspct	Unmarried	White	Male
18	Sales	Husband	White	Male
19	Exec-managerial	Unmarried	White	Female
20	Prof-specialty	Husband	White	Male
21	Other-service	Unmarried	Black	Female
22	Farming-fishing	Husband	Black	Male
23	Transport-moving	Husband	White	Male
24	Tech-support	Unmarried	White	Female
25	Tech-support	Husband	White	Male
26	Craft-repair	Own-child	White	Male
27	?	Husband	Asian-Pac-Islander	Male
28	Exec-managerial	Not-in-family	White	Male
29	Craft-repair	Husband	White	Male
...	...	...	...	...
32531	?	Not-in-family	Asian-Pac-Islander	Female
32532	Prof-specialty	Husband	White	Male
32533	Exec-managerial	Husband	Asian-Pac-Islander	Male
32534	Adm-clerical	Unmarried	White	Female
32535	Protective-serv	Own-child	Black	Male
32536	Exec-managerial	Not-in-family	White	Female
32537	Craft-repair	Not-in-family	Black	Male
32538	Prof-specialty	Unmarried	Black	Female
32539	?	Husband	White	Male
32540	Adm-clerical	Own-child	White	Female
32541	?	Not-in-family	Black	Female
32542	?	Husband	White	Male
32543	Prof-specialty	Unmarried	White	Female
32544	Other-service	Not-in-family	Other	Female
32545	Adm-clerical	Wife	White	Female
32546	Tech-support	Not-in-family	White	Female
32547	Machine-op-inspct	Husband	White	Male
32548	Prof-specialty	Not-in-family	White	Male
32549	Adm-clerical	Other-relative	White	Female
32550	Craft-repair	Husband	White	Male
32551	Handlers-cleaners	Husband	Amer-Indian-Eskimo	Male
32552	Sales	Husband	White	Male
32553	Tech-support	Not-in-family	Asian-Pac-Islander	Male
32554	Exec-managerial	Husband	White	Male
32555	Protective-serv	Not-in-family	White	Male
32556	Tech-support	Wife	White	Female
32557	Machine-op-inspct	Husband	White	Male
32558	Adm-clerical	Unmarried	White	Female
32559	Adm-clerical	Own-child	White	Male
32560	Exec-managerial	Wife	White	Female

```
              nc_num
0      United-States
1      United-States
2      United-States
3      United-States
4               Cuba
5      United-States
6            Jamaica
7      United-States
8      United-States
9      United-States
10     United-States
11             India
12     United-States
13     United-States
14                 ?
15            Mexico
16     United-States
17     United-States
18     United-States
19     United-States
20     United-States
21     United-States
22     United-States
23     United-States
24     United-States
25     United-States
26     United-States
27             South
28     United-States
29     United-States
...               ...
32531  United-States
32532  United-States
32533          Japan
32534  United-States
32535  United-States
32536  United-States
32537  United-States
32538  United-States
32539  United-States
32540  United-States
32541  United-States
32542  United-States
32543  United-States
32544  United-States
32545  United-States
```

```
32546    United-States
32547          Mexico
32548    United-States
32549    United-States
32550    United-States
32551    United-States
32552    United-States
32553          Taiwan
32554    United-States
32555    United-States
32556    United-States
32557    United-States
32558    United-States
32559    United-States
32560    United-States

[32561 rows x 22 columns]
```

In [34]: dummy = [1 if '-gov' in adult.workclass_num else 0 for adult.workclass_num in adult]

In [35]: dummy

Out[35]: [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

In [ ]: #workclass_num 1=government employee; 0=non government employee or out of work
        #ms_num 1=married; 0= single
        #rel_num 1=husband or wife; 0=not-in-family, unmarried, other-relative, own-child
        #race_num 1=white; 0=black, asian-pacislander, amer-indian-eskimo, other
        #sex_num 1=female; 0=male
        #nc_num 1=united state; 0= all other countries and unknowns
        #edu_num 1="years/grades" 13 to 16; 0="years/grades" 1 to 12
        #salary_num 1= >=50K; 0= <=50K

In [49]: adult = pd.read_csv('../data/adult_data_dummy.csv')
         adult

Out[49]:   age   workclass fnlwgt   education education_num \
         0   62   Private 416829   11th   7
         1   25   Private 197728   Some-college   10
         2   47   Self-emp-not-inc 229394   11th   7
         3   36   Private 177285   Assoc-voc   11
         4   47   Private   97419   HS-grad   9
         5   57   Private 173580   Some-college   10
         6   32   Federal-gov 504951   Assoc-acdm   12
         7   31   Private 288419   10th   6
         8   39   Private 154641   Assoc-voc   11
         9   39   Private 147258   10th   6
         10   47   Private   34458   HS-grad   9
         11   34   Private 473133   HS-grad   9
```

| 12 | 33 | Private | 226267 | 7th-8th | 4 |
|---|---|---|---|---|---|
| 13 | 17 | Private | 270942 | 5th-6th | 3 |
| 14 | 34 | Private | 133503 | Some-college | 10 |
| 15 | 24 | Private | 258700 | 5th-6th | 3 |
| 16 | 30 | Private | 54334 | 9th | 5 |
| 17 | 58 | ? | 114362 | Some-college | 10 |
| 18 | 39 | Private | 225504 | Masters | 14 |
| 19 | 23 | Self-emp-not-inc | 409230 | 1st-4th | 2 |
| 20 | 22 | Private | 361138 | HS-grad | 9 |
| 21 | 18 | Private | 127273 | 11th | 7 |
| 22 | 57 | Private | 99364 | 9th | 5 |
| 23 | 18 | Private | 145005 | Some-college | 10 |
| 24 | 31 | Self-emp-inc | 304212 | Some-college | 10 |
| 25 | 32 | Private | 42485 | HS-grad | 9 |
| 26 | 53 | Private | 49715 | HS-grad | 9 |
| 27 | 23 | Private | 73514 | HS-grad | 9 |
| 28 | 31 | Private | 217803 | HS-grad | 9 |
| 29 | 30 | Private | 271933 | HS-grad | 9 |
| ... | ... | ... | ... | ... | ... |
| 32531 | 39 | Self-emp-inc | 172538 | Bachelors | 13 |
| 32532 | 49 | Self-emp-not-inc | 102110 | HS-grad | 9 |
| 32533 | 29 | Private | 74500 | Bachelors | 13 |
| 32534 | 26 | Private | 222539 | Some-college | 10 |
| 32535 | 50 | Private | 48358 | HS-grad | 9 |
| 32536 | 41 | Private | 153031 | Some-college | 10 |
| 32537 | 42 | Private | 78765 | Some-college | 10 |
| 32538 | 48 | Private | 212954 | Some-college | 10 |
| 32539 | 36 | Private | 68273 | Bachelors | 13 |
| 32540 | 52 | Private | 114971 | Some-college | 10 |
| 32541 | 29 | Private | 100764 | Bachelors | 13 |
| 32542 | 68 | Private | 218637 | Some-college | 10 |
| 32543 | 52 | Private | 162238 | Bachelors | 13 |
| 32544 | 46 | Self-emp-inc | 285335 | Some-college | 10 |
| 32545 | 37 | Private | 38948 | HS-grad | 9 |
| 32546 | 40 | Private | 209547 | Bachelors | 13 |
| 32547 | 57 | Self-emp-not-inc | 225913 | Some-college | 10 |
| 32548 | 34 | Private | 24266 | Assoc-voc | 11 |
| 32549 | 48 | Self-emp-not-inc | 107231 | Bachelors | 13 |
| 32550 | 40 | Private | 111067 | Bachelors | 13 |
| 32551 | 67 | Self-emp-inc | 182581 | Some-college | 10 |
| 32552 | 39 | Federal-gov | 184964 | Some-college | 10 |
| 32553 | 47 | Private | 30840 | Masters | 14 |
| 32554 | 40 | Private | 356934 | Assoc-acdm | 12 |
| 32555 | 44 | Private | 157249 | Bachelors | 13 |
| 32556 | 38 | Self-emp-not-inc | 164526 | Prof-school | 15 |
| 32557 | 52 | Private | 110977 | Doctorate | 16 |
| 32558 | 48 | Federal-gov | 33109 | Bachelors | 13 |
| 32559 | 50 | Private | 100109 | Bachelors | 13 |

```
32560   57         Private  186386      Bachelors          13

          marital_status           occupation   relationship  \
0              Separated         Other-service  Not-in-family
1          Never-married          Adm-clerical      Own-child
2               Divorced       Exec-managerial      Unmarried
3          Never-married         Prof-specialty      Unmarried
4     Married-civ-spouse       Protective-serv           Wife
5                Widowed      Machine-op-inspct      Unmarried
6     Married-civ-spouse       Protective-serv        Husband
7     Married-civ-spouse      Machine-op-inspct        Husband
8     Married-civ-spouse          Adm-clerical        Husband
9     Married-civ-spouse          Craft-repair        Husband
10    Married-civ-spouse       Farming-fishing        Husband
11    Married-civ-spouse                 Sales        Husband
12         Never-married                 Sales  Not-in-family
13         Never-married         Other-service  Other-relative
14              Divorced      Transport-moving  Not-in-family
15         Never-married       Farming-fishing  Other-relative
16         Never-married                 Sales  Not-in-family
17    Married-civ-spouse                     ?        Husband
18         Never-married         Prof-specialty  Not-in-family
19    Married-civ-spouse                 Sales  Other-relative
20         Never-married     Handlers-cleaners  Other-relative
21         Never-married         Other-service  Other-relative
22    Married-civ-spouse      Machine-op-inspct        Husband
23         Never-married         Other-service      Own-child
24         Never-married       Exec-managerial      Own-child
25              Divorced          Craft-repair      Unmarried
26              Divorced      Transport-moving      Unmarried
27         Never-married          Adm-clerical      Own-child
28              Divorced                 Sales  Not-in-family
29              Divorced          Adm-clerical  Not-in-family
...                  ...                   ...            ...
32531 Married-civ-spouse         Prof-specialty        Husband
32532 Married-civ-spouse          Tech-support        Husband
32533 Married-civ-spouse                 Sales        Husband
32534 Married-civ-spouse          Craft-repair        Husband
32535 Married-civ-spouse          Craft-repair        Husband
32536 Married-civ-spouse                 Sales        Husband
32537 Married-civ-spouse          Craft-repair        Husband
32538 Married-civ-spouse          Craft-repair        Husband
32539 Married-civ-spouse                 Sales        Husband
32540 Married-civ-spouse          Craft-repair        Husband
32541 Married-civ-spouse       Exec-managerial        Husband
32542 Married-civ-spouse                 Sales        Husband
32543 Married-civ-spouse         Prof-specialty        Husband
32544 Married-civ-spouse                 Sales        Husband
```

```
32545  Married-civ-spouse       Prof-specialty         Husband
32546  Married-civ-spouse       Prof-specialty         Husband
32547  Married-civ-spouse    Transport-moving          Husband
32548  Married-civ-spouse          Craft-repair        Husband
32549  Married-civ-spouse     Exec-managerial          Husband
32550  Married-civ-spouse   Handlers-cleaners          Husband
32551  Married-civ-spouse     Exec-managerial          Husband
32552            Divorced        Adm-clerical      Not-in-family
32553            Divorced      Prof-specialty      Not-in-family
32554            Divorced     Exec-managerial      Not-in-family
32555  Married-civ-spouse      Prof-specialty          Husband
32556       Never-married      Prof-specialty      Not-in-family
32557       Never-married     Exec-managerial      Not-in-family
32558            Divorced     Exec-managerial        Unmarried
32559            Divorced     Exec-managerial        Unmarried
32560             Widowed      Prof-specialty        Unmarried


                     race      sex    ...    native_country   salary_50K  \
0                    Black   Female    ...     United-States      <=50K
1                    White   Female    ...     United-States      <=50K
2                    White   Female    ...     United-States      <=50K
3                    Black   Female    ...     United-States      <=50K
4                    Black   Female    ...     United-States      <=50K
5                    White   Female    ...     United-States      <=50K
6                    White     Male    ...     United-States      <=50K
7                    White     Male    ...     United-States      <=50K
8                    White     Male    ...     United-States      <=50K
9                    White     Male    ...     United-States      <=50K
10                   White     Male    ...     United-States      <=50K
11                   White     Male    ...     United-States      <=50K
12                   White     Male    ...            Mexico      <=50K
13                   White     Male    ...            Mexico      <=50K
14                   White     Male    ...     United-States      <=50K
15                   Black     Male    ...            Mexico      <=50K
16                   White     Male    ...     United-States      <=50K
17                   Black     Male    ...     United-States      <=50K
18                   White     Male    ...     United-States      <=50K
19                   White     Male    ...     United-States      <=50K
20                   White     Male    ...     United-States      <=50K
21                   White     Male    ...     United-States      <=50K
22                   White     Male    ...       Puerto-Rico      <=50K
23                   White     Male    ...     United-States      <=50K
24                   White     Male    ...     United-States      <=50K
25                   White     Male    ...     United-States      <=50K
26                   White     Male    ...     United-States      <=50K
27      Asian-Pac-Islander   Female    ...           Vietnam      <=50K
28                   White   Female    ...     United-States      <=50K
29                   White   Female    ...     United-States      <=50K
```

```
...           ...    ...   ...                    ...        ...
32531         White   Male  ...        United-States        >50K
32532         White   Male  ...        United-States        >50K
32533         White   Male  ...        United-States        >50K
32534         White   Male  ...        United-States        >50K
32535         White   Male  ...        United-States        >50K
32536         White   Male  ...        United-States        >50K
32537         White   Male  ...        United-States        >50K
32538         White   Male  ...        United-States        >50K
32539         White   Male  ...        United-States        >50K
32540         White   Male  ...        United-States        >50K
32541         White   Male  ...        United-States        >50K
32542         White   Male  ...        United-States        >50K
32543         White   Male  ...        United-States        >50K
32544         White   Male  ...        United-States        >50K
32545         White   Male  ...        United-States        >50K
32546         White   Male  ...        United-States        >50K
32547         White   Male  ...        United-States        >50K
32548         White   Male  ...        United-States        >50K
32549         White   Male  ...        United-States        >50K
32550         White   Male  ...        United-States        >50K
32551         White   Male  ...        United-States        >50K
32552         White   Male  ...        United-States        >50K
32553         White   Male  ...        United-States        >50K
32554         White   Male  ...        United-States        >50K
32555         Black   Male  ...        United-States        >50K
32556         White   Male  ...        United-States        >50K
32557         White   Male  ...        United-States        >50K
32558         White   Male  ...        United-States        >50K
32559         White   Male  ...        United-States        >50K
32560         White   Male  ...        United-States        >50K

      workclass_num ms_num rel_num race_num sex_num nc_num edu_num  \
0                 0      0       0        0       1      1       0
1                 0      0       0        1       1      1       0
2                 0      0       0        1       1      1       0
3                 0      0       0        0       1      1       0
4                 0      1       1        0       1      1       0
5                 0      0       0        1       1      1       0
6                 1      1       1        1       0      1       0
7                 0      1       1        1       0      1       0
8                 0      1       1        1       0      1       0
9                 0      1       1        1       0      1       0
10                0      1       1        1       0      1       0
11                0      1       1        1       0      1       0
12                0      0       0        1       0      0       0
13                0      0       0        1       0      0       0
14                0      0       0        1       0      1       0
```

15	0	0	0	0	0	0	0
16	0	0	0	1	0	1	0
17	0	1	1	0	0	1	0
18	0	0	0	1	0	1	0
19	0	1	0	1	0	1	0
20	0	0	0	1	0	1	0
21	0	0	0	1	0	1	0
22	0	1	1	1	0	0	0
23	0	0	0	1	0	1	0
24	0	0	0	1	0	1	0
25	0	0	0	1	0	1	0
26	0	0	0	1	0	1	0
27	0	0	0	0	1	0	0
28	0	0	0	1	1	1	0
29	0	0	0	1	1	1	0
...	...	...	...	...	...	...	...
32531	0	1	1	1	0	1	1
32532	0	1	1	1	0	1	1
32533	0	1	1	1	0	1	1
32534	0	1	1	1	0	1	1
32535	0	1	1	1	0	1	1
32536	0	1	1	1	0	1	1
32537	0	1	1	1	0	1	1
32538	0	1	1	1	0	1	1
32539	0	1	1	1	0	1	1
32540	0	1	1	1	0	1	1
32541	0	1	1	1	0	1	1
32542	0	1	1	1	0	1	1
32543	0	1	1	1	0	1	1
32544	0	1	1	1	0	1	1
32545	0	1	1	1	0	1	1
32546	0	1	1	1	0	1	1
32547	0	1	1	1	0	1	1
32548	0	1	1	1	0	1	1
32549	0	1	1	1	0	1	1
32550	0	1	1	1	0	1	1
32551	0	1	1	1	0	1	1
32552	1	0	0	1	0	1	1
32553	0	0	0	1	0	1	1
32554	0	0	0	1	0	1	1
32555	0	1	1	0	0	1	1
32556	0	0	0	1	0	1	1
32557	0	0	0	1	0	1	1
32558	1	0	0	1	0	1	1
32559	0	0	0	1	0	1	1
32560	0	0	0	1	0	1	1

salary_num

```
0           0
1           0
2           0
3           0
4           0
5           0
6           0
7           0
8           0
9           0
10          0
11          0
12          0
13          0
14          0
15          0
16          0
17          0
18          0
19          0
20          0
21          0
22          0
23          0
24          0
25          0
26          0
27          0
28          0
29          0
...         ...
32531       1
32532       1
32533       1
32534       1
32535       1
32536       1
32537       1
32538       1
32539       1
32540       1
32541       1
32542       1
32543       1
32544       1
32545       1
32546       1
32547       1
```

```
32548          1
32549          1
32550          1
32551          1
32552          1
32553          1
32554          1
32555          1
32556          1
32557          1
32558          1
32559          1
32560          1

[32561 rows x 23 columns]
```

In [50]: adult.edu_num.value_counts(normalize=True)

Out[50]: 0   0.75225
         1   0.24775
         Name: edu_num, dtype: float64

In [51]: adult.edu_num.value_counts(normalize=False)

Out[51]: 0   24494
         1   8067
         Name: edu_num, dtype: int64

In [54]: adult.salary_num.value_counts(normalize=True)

Out[54]: 0   0.75919
         1   0.24081
         Name: salary_num, dtype: float64

In [55]: adult.salary_num.value_counts(normalize=False)

Out[55]: 0   24720
         1   7841
         Name: salary_num, dtype: int64

In [57]: fig, ax = plt.subplots()
         adult[adult.salary_num==0].plot(kind='scatter', x='education_num', y='age', color='blu
         adult[adult.salary_num==1].plot(kind='scatter', x='education_num', y='age', color='gre
         # plt.scatter(adult.education_num, adult.age, alpha=0.2)

Out[57]: <matplotlib.axes._subplots.AxesSubplot at 0x1a6e1992828>
```
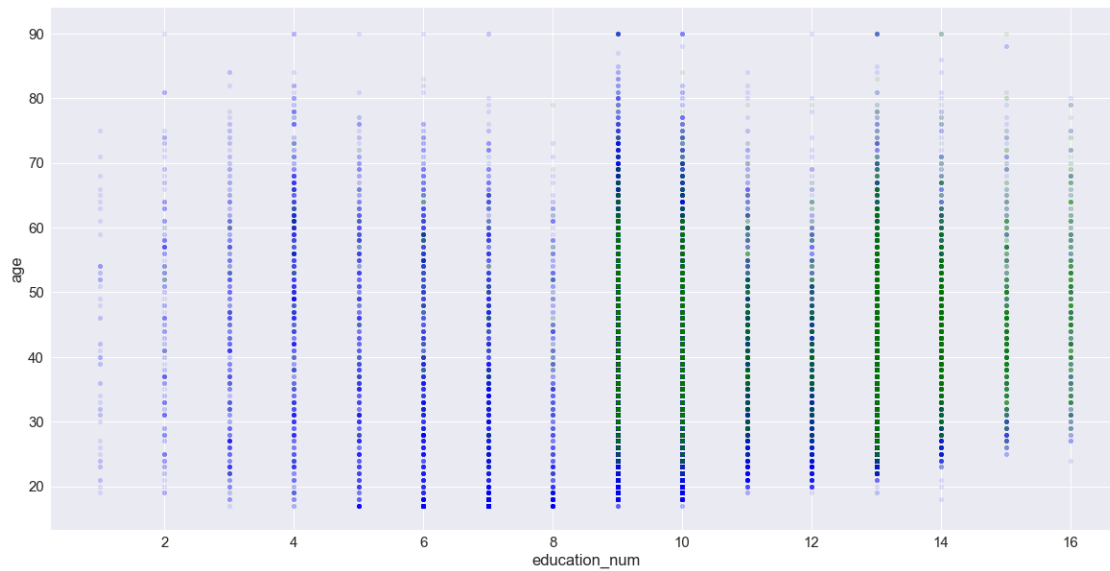
```
In [64]: adult[adult.salary_num==0].education_num.hist(bins=np.arange(0,30,1), alpha=.4)
         adult[adult.salary_num==1].education_num.hist(bins=np.arange(0,30,1), alpha=.4)
```

```
Out[64]: <matplotlib.axes._subplots.AxesSubplot at 0x1a6e4626e10>
```
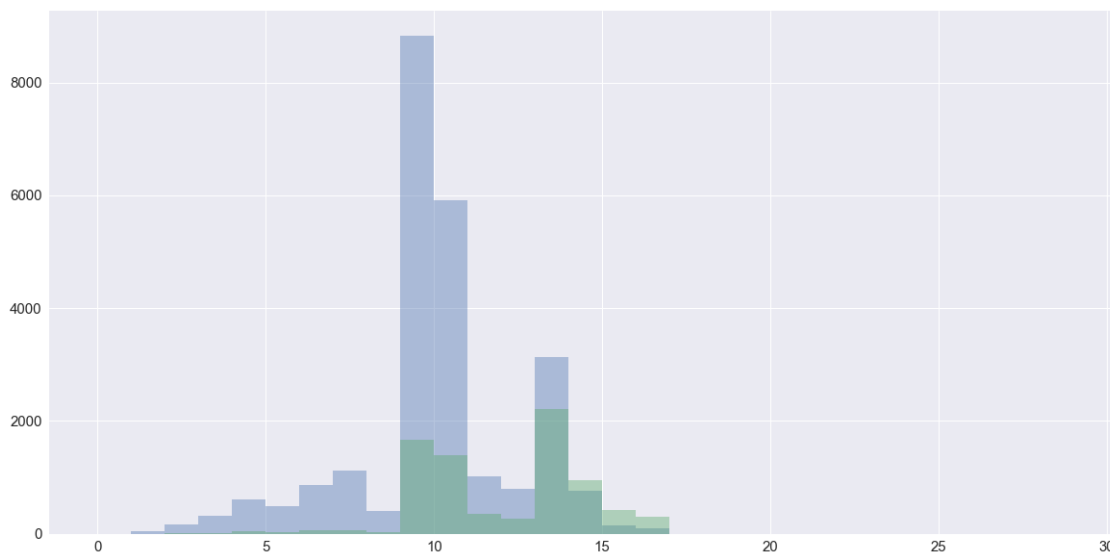


```
In [59]: fig, ax = plt.subplots()
         adult[adult.salary_num==0].plot(kind='scatter', x='race_num', y='age', color='blue', a
         adult[adult.salary_num==1].plot(kind='scatter', x='race_num', y='age', color='green',
         # plt.scatter(adult.race_num, adult.age, alpha=0.2)
```
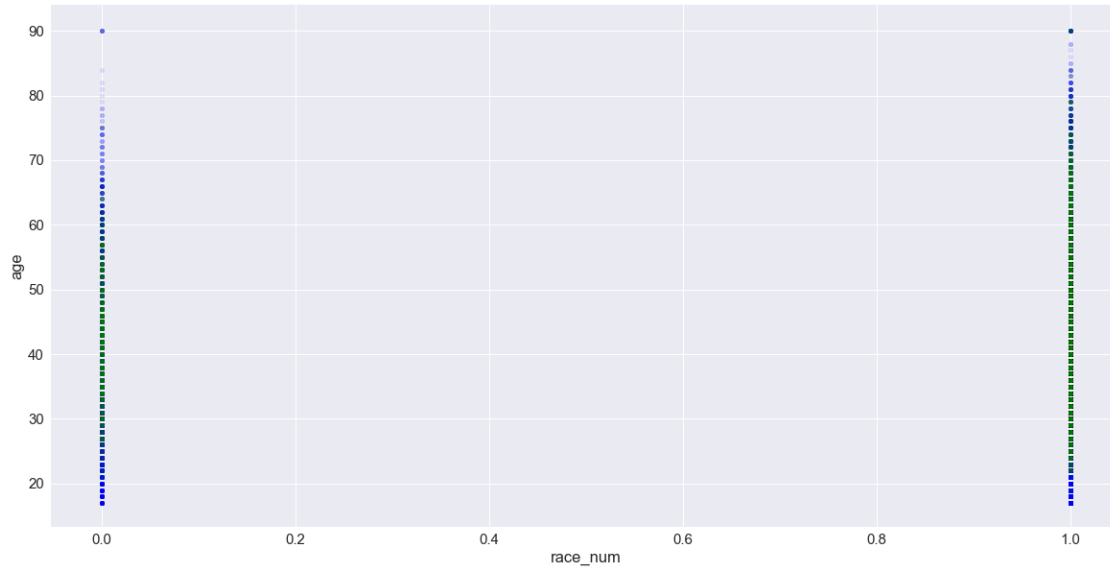
`<matplotlib.axes._subplots.AxesSubplot at 0x1a6e33edb00>`

In [60]:
```python
fig, ax = plt.subplots()
adult[adult.salary_num==0].plot(kind='scatter', x='workclass_num', y='age', color='blu
adult[adult.salary_num==1].plot(kind='scatter', x='workclass_num', y='age', color='gre
# plt.scatter(adult.workclass_num, adult.age, alpha=0.2)
```

Out[60]: `<matplotlib.axes._subplots.AxesSubplot at 0x1a6e229cc50>`
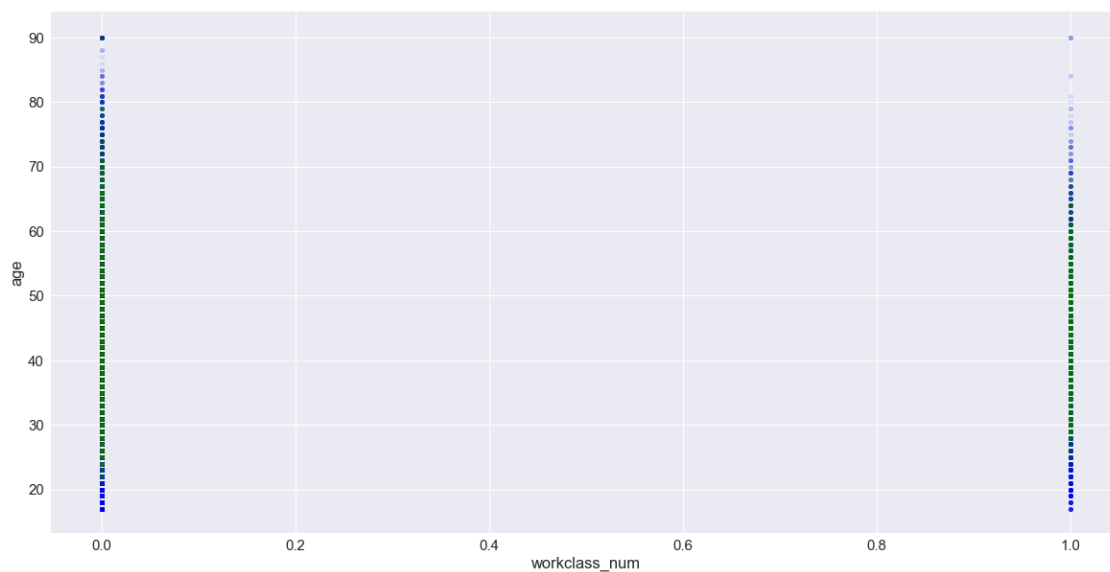
```
In [65]: # fit a linear regression model and store the predictions
         feature_cols = ['salary_num', 'edu_num', 'workclass_num', 'sex_num', 'rel_num', 'ms_n
         X = adult[feature_cols]
         y = adult.age

         model = Model()
         model.fit(X, y)
         adult['pred'] = model.predict(X)


         from sklearn.metrics import accuracy_score
         accuracy_score(adult.age, adult.pred.round())

Out[65]: 0.044347532323945824

In [72]: adult
```

Out[72]:

	age	workclass	fnlwgt	education	education_num \
0	62	Private	416829	11th	7
1	25	Private	197728	Some-college	10
2	47	Self-emp-not-inc	229394	11th	7
3	36	Private	177285	Assoc-voc	11
4	47	Private	97419	HS-grad	9
5	57	Private	173580	Some-college	10
6	32	Federal-gov	504951	Assoc-acdm	12
7	31	Private	288419	10th	6
8	39	Private	154641	Assoc-voc	11
9	39	Private	147258	10th	6
10	47	Private	34458	HS-grad	9
11	34	Private	473133	HS-grad	9
12	33	Private	226267	7th-8th	4
13	17	Private	270942	5th-6th	3
14	34	Private	133503	Some-college	10
15	24	Private	258700	5th-6th	3
16	30	Private	54334	9th	5
17	58	?	114362	Some-college	10
18	39	Private	225504	Masters	14
19	23	Self-emp-not-inc	409230	1st-4th	2
20	22	Private	361138	HS-grad	9
21	18	Private	127273	11th	7
22	57	Private	99364	9th	5
23	18	Private	145005	Some-college	10
24	31	Self-emp-inc	304212	Some-college	10
25	32	Private	42485	HS-grad	9
26	53	Private	49715	HS-grad	9
27	23	Private	73514	HS-grad	9
28	31	Private	217803	HS-grad	9
29	30	Private	271933	HS-grad	9

```
 ...    ...                 ...     ...              ...            ...
32531   39        Self-emp-inc  172538       Bachelors           13
32532   49    Self-emp-not-inc  102110         HS-grad            9
32533   29             Private   74500       Bachelors           13
32534   26             Private  222539    Some-college           10
32535   50             Private   48358         HS-grad            9
32536   41             Private  153031    Some-college           10
32537   42             Private   78765    Some-college           10
32538   48             Private  212954    Some-college           10
32539   36             Private   68273       Bachelors           13
32540   52             Private  114971    Some-college           10
32541   29             Private  100764       Bachelors           13
32542   68             Private  218637    Some-college           10
32543   52             Private  162238       Bachelors           13
32544   46        Self-emp-inc  285335    Some-college           10
32545   37             Private   38948         HS-grad            9
32546   40             Private  209547       Bachelors           13
32547   57    Self-emp-not-inc  225913    Some-college           10
32548   34             Private   24266        Assoc-voc           11
32549   48    Self-emp-not-inc  107231       Bachelors           13
32550   40             Private  111067       Bachelors           13
32551   67        Self-emp-inc  182581    Some-college           10
32552   39         Federal-gov  184964    Some-college           10
32553   47             Private   30840         Masters           14
32554   40             Private  356934       Assoc-acdm          12
32555   44             Private  157249       Bachelors           13
32556   38    Self-emp-not-inc  164526     Prof-school           15
32557   52             Private  110977        Doctorate          16
32558   48         Federal-gov   33109       Bachelors           13
32559   50             Private  100109       Bachelors           13
32560   57             Private  186386       Bachelors           13

          marital_status          occupation      relationship  \
0              Separated       Other-service    Not-in-family
1          Never-married        Adm-clerical        Own-child
2               Divorced     Exec-managerial        Unmarried
3          Never-married       Prof-specialty        Unmarried
4     Married-civ-spouse     Protective-serv             Wife
5                Widowed    Machine-op-inspct        Unmarried
6     Married-civ-spouse     Protective-serv          Husband
7     Married-civ-spouse    Machine-op-inspct          Husband
8     Married-civ-spouse        Adm-clerical          Husband
9     Married-civ-spouse         Craft-repair          Husband
10    Married-civ-spouse      Farming-fishing          Husband
11    Married-civ-spouse               Sales          Husband
12         Never-married               Sales    Not-in-family
13         Never-married       Other-service   Other-relative
14              Divorced    Transport-moving    Not-in-family
```

15	Never-married	Farming-fishing	Other-relative
16	Never-married	Sales	Not-in-family
17	Married-civ-spouse	?	Husband
18	Never-married	Prof-specialty	Not-in-family
19	Married-civ-spouse	Sales	Other-relative
20	Never-married	Handlers-cleaners	Other-relative
21	Never-married	Other-service	Other-relative
22	Married-civ-spouse	Machine-op-inspct	Husband
23	Never-married	Other-service	Own-child
24	Never-married	Exec-managerial	Own-child
25	Divorced	Craft-repair	Unmarried
26	Divorced	Transport-moving	Unmarried
27	Never-married	Adm-clerical	Own-child
28	Divorced	Sales	Not-in-family
29	Divorced	Adm-clerical	Not-in-family
...	...	...	...
32531	Married-civ-spouse	Prof-specialty	Husband
32532	Married-civ-spouse	Tech-support	Husband
32533	Married-civ-spouse	Sales	Husband
32534	Married-civ-spouse	Craft-repair	Husband
32535	Married-civ-spouse	Craft-repair	Husband
32536	Married-civ-spouse	Sales	Husband
32537	Married-civ-spouse	Craft-repair	Husband
32538	Married-civ-spouse	Craft-repair	Husband
32539	Married-civ-spouse	Sales	Husband
32540	Married-civ-spouse	Craft-repair	Husband
32541	Married-civ-spouse	Exec-managerial	Husband
32542	Married-civ-spouse	Sales	Husband
32543	Married-civ-spouse	Prof-specialty	Husband
32544	Married-civ-spouse	Sales	Husband
32545	Married-civ-spouse	Prof-specialty	Husband
32546	Married-civ-spouse	Prof-specialty	Husband
32547	Married-civ-spouse	Transport-moving	Husband
32548	Married-civ-spouse	Craft-repair	Husband
32549	Married-civ-spouse	Exec-managerial	Husband
32550	Married-civ-spouse	Handlers-cleaners	Husband
32551	Married-civ-spouse	Exec-managerial	Husband
32552	Divorced	Adm-clerical	Not-in-family
32553	Divorced	Prof-specialty	Not-in-family
32554	Divorced	Exec-managerial	Not-in-family
32555	Married-civ-spouse	Prof-specialty	Husband
32556	Never-married	Prof-specialty	Not-in-family
32557	Never-married	Exec-managerial	Not-in-family
32558	Divorced	Exec-managerial	Unmarried
32559	Divorced	Exec-managerial	Unmarried
32560	Widowed	Prof-specialty	Unmarried

```
                          race      sex  ...   salary_50K  workclass_num  ms_num  \
```

0	Black	Female	...	<=50K	0	0
1	White	Female	...	<=50K	0	0
2	White	Female	...	<=50K	0	0
3	Black	Female	...	<=50K	0	0
4	Black	Female	...	<=50K	0	1
5	White	Female	...	<=50K	0	0
6	White	Male	...	<=50K	1	1
7	White	Male	...	<=50K	0	1
8	White	Male	...	<=50K	0	1
9	White	Male	...	<=50K	0	1
10	White	Male	...	<=50K	0	1
11	White	Male	...	<=50K	0	1
12	White	Male	...	<=50K	0	0
13	White	Male	...	<=50K	0	0
14	White	Male	...	<=50K	0	0
15	Black	Male	...	<=50K	0	0
16	White	Male	...	<=50K	0	0
17	Black	Male	...	<=50K	0	1
18	White	Male	...	<=50K	0	0
19	White	Male	...	<=50K	0	1
20	White	Male	...	<=50K	0	0
21	White	Male	...	<=50K	0	0
22	White	Male	...	<=50K	0	1
23	White	Male	...	<=50K	0	0
24	White	Male	...	<=50K	0	0
25	White	Male	...	<=50K	0	0
26	White	Male	...	<=50K	0	0
27	Asian-Pac-Islander	Female	...	<=50K	0	0
28	White	Female	...	<=50K	0	0
29	White	Female	...	<=50K	0	0
...	...	...	...	...	...	...
32531	White	Male	...	>50K	0	1
32532	White	Male	...	>50K	0	1
32533	White	Male	...	>50K	0	1
32534	White	Male	...	>50K	0	1
32535	White	Male	...	>50K	0	1
32536	White	Male	...	>50K	0	1
32537	White	Male	...	>50K	0	1
32538	White	Male	...	>50K	0	1
32539	White	Male	...	>50K	0	1
32540	White	Male	...	>50K	0	1
32541	White	Male	...	>50K	0	1
32542	White	Male	...	>50K	0	1
32543	White	Male	...	>50K	0	1
32544	White	Male	...	>50K	0	1
32545	White	Male	...	>50K	0	1
32546	White	Male	...	>50K	0	1
32547	White	Male	...	>50K	0	1

```
32548              White    Male  ...      >50K          0      1
32549              White    Male  ...      >50K          0      1
32550              White    Male  ...      >50K          0      1
32551              White    Male  ...      >50K          0      1
32552              White    Male  ...      >50K          1      0
32553              White    Male  ...      >50K          0      0
32554              White    Male  ...      >50K          0      0
32555              Black    Male  ...      >50K          0      1
32556              White    Male  ...      >50K          0      0
32557              White    Male  ...      >50K          0      0
32558              White    Male  ...      >50K          1      0
32559              White    Male  ...      >50K          0      0
32560              White    Male  ...      >50K          0      0

       rel_num race_num  sex_num  nc_num  edu_num  salary_num  pred
0            0        0        1       1        0           0    20
1            0        1        1       1        0           0    20
2            0        1        1       1        0           0    20
3            0        0        1       1        0           0    20
4            1        0        1       1        0           0    36
5            0        1        1       1        0           0    20
6            1        1        0       1        0           0    43
7            1        1        0       1        0           0    35
8            1        1        0       1        0           0    35
9            1        1        0       1        0           0    35
10           1        1        0       1        0           0    35
11           1        1        0       1        0           0    35
12           0        1        0       0        0           0    23
13           0        1        0       0        0           0    23
14           0        1        0       1        0           0    23
15           0        0        0       0        0           0    23
16           0        1        0       1        0           0    23
17           1        0        0       1        0           0    35
18           0        1        0       1        0           0    23
19           0        1        0       1        0           0    25
20           0        1        0       1        0           0    23
21           0        1        0       1        0           0    23
22           1        1        0       0        0           0    33
23           0        1        0       1        0           0    23
24           0        1        0       1        0           0    23
25           0        1        0       1        0           0    23
26           0        1        0       1        0           0    23
27           0        0        1       0        0           0    23
28           0        1        1       1        0           0    20
29           0        1        1       1        0           0    20
...        ...      ...      ...     ...      ...         ...   ...
32531        1        1        0       1        1           1    38
32532        1        1        0       1        1           1    38
```

```
32533        1        1        0        1        1        1     38
32534        1        1        0        1        1        1     38
32535        1        1        0        1        1        1     38
32536        1        1        0        1        1        1     38
32537        1        1        0        1        1        1     38
32538        1        1        0        1        1        1     38
32539        1        1        0        1        1        1     38
32540        1        1        0        1        1        1     38
32541        1        1        0        1        1        1     38
32542        1        1        0        1        1        1     38
32543        1        1        0        1        1        1     38
32544        1        1        0        1        1        1     38
32545        1        1        0        1        1        1     38
32546        1        1        0        1        1        1     38
32547        1        1        0        1        1        1     38
32548        1        1        0        1        1        1     38
32549        1        1        0        1        1        1     38
32550        1        1        0        1        1        1     38
32551        1        1        0        1        1        1     38
32552        0        1        0        1        1        1     45
32553        0        1        0        1        1        1     37
32554        0        1        0        1        1        1     37
32555        1        0        0        1        1        1     38
32556        0        1        0        1        1        1     37
32557        0        1        0        1        1        1     37
32558        0        1        0        1        1        1     45
32559        0        1        0        1        1        1     37
32560        0        1        0        1        1        1     37

[32561 rows x 24 columns]
```

In [67]: adult[adult.salary_num==0].age.hist(bins=np.arange(0,100,1), alpha=.4)
         adult[adult.salary_num==1].age.hist(bins=np.arange(0,100,1), alpha=.4)

Out[67]: <matplotlib.axes._subplots.AxesSubplot at 0x1a6e59b7080>
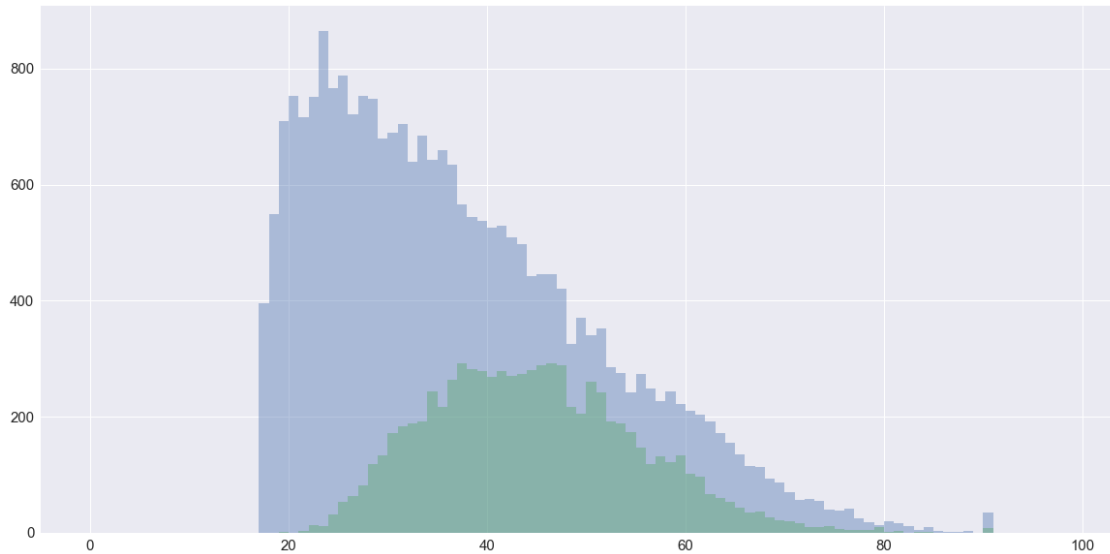
```
In [77]: adult[(adult.pred==0) & adult[adult.salary_num==0].age.hist(bins=np.arange(0,100,1),
             adult[(adult.pred==20) & adult[adult.salary_num==1].age.hist(bins=np.arange(0,100,1),


           File "<ipython-input-77-b709bebf9846>", line 2
         adult[(adult.pred==20) & adult[adult.salary_num==1].age.hist(bins=np.arange(0,100,1),
               ^
       SyntaxError: invalid syntax
```

```
In [81]: # fit a linear regression model and store the predictions
         feature_cols = ['salary_num', 'edu_num', 'workclass_num', 'sex_num', 'rel_num', 'ms_nu
         X = adult[feature_cols]
         y = adult.age

         model = Model()
         model.fit(X, y)
         adult['pred'] = model.predict(X)


         from sklearn.metrics import accuracy_score, confusion_matrix
         accuracy_score(adult.age, adult.pred.round())
         confusion_matrix(adult.age, adult.pred.round())

Out[81]: array([[  0,   0, 114, ...,   0,   0,   0],
                [  0,   0, 141, ...,   0,   0,   0],
                [  0,   0, 175, ...,   0,   0,   0],
                ...,
```

61

```
          [  0,    0,    0,  ...,    0,    0,    0],
          [  0,    0,    1,  ...,    0,    0,    0],
          [  0,    0,    7,  ...,    0,    0,    0]], dtype=int64)
```

In [82]: [[num of correct miss pred] [num of inc miss pred]]


```
      File "<ipython-input-82-8531411831fc>", line 1
    [[num of correct miss pred] [num of inc miss pred]]
          ^
SyntaxError: invalid syntax
```


In [88]: `from sklearn.metrics import classification_report`
         `print(classification_report(adult.age, adult.pred.round()))`

	precision	recall	f1-score	support
17	0.00	0.00	0.00	395
18	0.00	0.00	0.00	550
19	0.06	0.25	0.10	712
20	0.04	0.28	0.08	753
21	0.00	0.00	0.00	720
22	0.00	0.00	0.00	765
23	0.05	0.48	0.09	877
24	0.00	0.00	0.00	798
25	0.06	0.01	0.02	841
26	0.00	0.00	0.00	785
27	0.00	0.00	0.00	835
28	0.05	0.02	0.02	867
29	0.00	0.00	0.00	813
30	0.04	0.01	0.02	861
31	0.04	0.06	0.05	888
32	0.00	0.00	0.00	828
33	0.03	0.02	0.02	875
34	0.00	0.00	0.00	886
35	0.04	0.18	0.06	876
36	0.03	0.02	0.02	898
37	0.04	0.15	0.06	858
38	0.04	0.07	0.05	827
39	0.04	0.03	0.03	816
40	0.02	0.00	0.00	794
41	0.03	0.02	0.03	808
42	0.00	0.00	0.00	780
43	0.04	0.05	0.04	770
44	0.00	0.00	0.00	724
45	0.04	0.01	0.02	734
46	0.05	0.06	0.06	737

47	0.05	0.06	0.05	708
48	0.00	0.00	0.00	543
49	0.00	0.00	0.00	577
50	0.00	0.00	0.00	602
51	0.00	0.00	0.00	595
52	0.00	0.00	0.00	478
53	0.00	0.00	0.00	464
54	0.00	0.00	0.00	415
55	0.00	0.00	0.00	419
56	0.00	0.00	0.00	366
57	0.00	0.00	0.00	358
58	0.00	0.00	0.00	366
59	0.00	0.00	0.00	355
60	0.00	0.00	0.00	312
61	0.00	0.00	0.00	300
62	0.00	0.00	0.00	258
63	0.00	0.00	0.00	230
64	0.00	0.00	0.00	208
65	0.00	0.00	0.00	178
66	0.00	0.00	0.00	150
67	0.00	0.00	0.00	151
68	0.00	0.00	0.00	120
69	0.00	0.00	0.00	108
70	0.00	0.00	0.00	89
71	0.00	0.00	0.00	72
72	0.00	0.00	0.00	67
73	0.00	0.00	0.00	64
74	0.00	0.00	0.00	51
75	0.00	0.00	0.00	45
76	0.00	0.00	0.00	46
77	0.00	0.00	0.00	29
78	0.00	0.00	0.00	23
79	0.00	0.00	0.00	22
80	0.00	0.00	0.00	22
81	0.00	0.00	0.00	20
82	0.00	0.00	0.00	12
83	0.00	0.00	0.00	6
84	0.00	0.00	0.00	10
85	0.00	0.00	0.00	3
86	0.00	0.00	0.00	1
87	0.00	0.00	0.00	1
88	0.00	0.00	0.00	3
90	0.00	0.00	0.00	43
avg / total	0.02	0.04	0.02	32561

```
C:\Users\Erin\Anaconda3\lib\site-packages\sklearn\metrics\classification.py:1135: UndefinedMet
  'precision', 'predicted', average, warn_for)


In [89]: model.coef_

Out[89]: array([[-2.49348757e+00,  3.53084329e-01, -1.12169513e+00,
                 -2.46632869e-01, -2.04133886e+00, -2.33650510e+00,
                  5.16716516e-01],
                [-2.76530317e+00,  3.35983392e-01, -1.22412472e+00,
                 -1.66556281e-01, -1.99139281e+00, -1.83833529e+00,
                  6.88736992e-01],
                [-2.57382149e+00,  3.09780780e-01, -9.73564952e-01,
                 -9.40877424e-02, -2.11335913e+00, -1.24742152e+00,
                  3.03289200e-01],
                [-3.16140407e+00, -1.85050830e-02, -8.89862599e-01,
                 -1.07520080e-01, -2.06906276e+00, -5.66202313e-01,
                  1.93215646e-01],
                [-2.70932997e+00,  6.53736918e-02, -8.29757610e-01,
                 -1.70297156e-01, -1.22465742e+00, -8.53598666e-01,
                  5.52388520e-02],
                [-1.90284765e+00,  2.94114042e-02, -4.78008878e-01,
                 -1.76133348e-01, -1.35026342e+00, -4.86710207e-01,
                 -5.62235661e-02],
                [-2.21426085e+00,  1.10770009e-01, -3.56314849e-01,
                 -3.20712793e-01, -1.07991148e+00, -5.68016255e-01,
                 -1.84372127e-01],
                [-1.60383023e+00,  3.63461501e-02, -2.03727445e-01,
                 -2.73706743e-01, -6.08753742e-01, -3.16314470e-01,
                 -1.33630372e-01],
                [-1.15306341e+00,  3.96335475e-02, -3.23908037e-01,
                 -2.68891524e-01, -8.49093329e-01,  4.17554639e-03,
                 -2.08965302e-01],
                [-1.07463082e+00, -2.43818576e-02, -1.80018406e-01,
                 -1.58725055e-01,  5.57932430e-02, -5.35032328e-01,
                 -1.44629178e-01],
                [-8.94474010e-01,  6.22708294e-02, -1.18706038e-01,
                 -2.19584903e-01, -3.68848280e-01, -8.94628750e-02,
                 -2.45633510e-01],
                [-5.92615866e-01, -9.73583648e-02, -1.97535636e-01,
                 -5.50215470e-02, -4.87223534e-01,  2.94974452e-01,
                 -2.65830003e-01],
                [-4.65703676e-01,  5.05485453e-02, -3.66094220e-02,
                 -1.02727244e-01, -3.36802452e-01,  2.76753745e-01,
                 -2.59788113e-01],
                [-2.81301733e-01,  7.37889718e-02, -1.05501526e-01,
                 -1.09822371e-01, -4.02521137e-01,  4.37975901e-01,
                 -2.37710385e-01],
```

```
[-2.71581368e-01,  7.77194134e-02, -2.98136210e-03,
 -7.64393423e-02,  3.31028934e-01, -2.44649063e-01,
 -1.53969717e-01],
[-1.93023847e-01,  1.74430161e-03, -3.71867089e-02,
 -4.68132494e-02, -7.34547609e-03,  2.32762452e-01,
 -1.11587956e-01],
[-2.35381274e-01, -1.80471518e-02,  2.70789542e-02,
 -1.78820508e-01, -1.02337504e-01,  2.34381468e-01,
 -1.37574453e-01],
[ 5.01903132e-02,  1.50897861e-02, -4.77826168e-02,
 -8.71782032e-02,  1.77431316e-01,  4.10681141e-02,
 -1.47020536e-01],
[-1.14849723e-01, -9.24717058e-02, -2.72929474e-01,
 -1.38096996e-01,  5.72958581e-02,  2.07808268e-01,
  4.21380174e-02],
[ 1.60297090e-01, -2.10928972e-01,  5.08514398e-02,
  1.05147167e-01,  5.12365019e-01, -2.53327639e-01,
 -1.01365198e-01],
[ 3.07386182e-01, -1.99400545e-02,  3.29130398e-02,
 -9.32418122e-02,  4.64231771e-01, -1.33943170e-01,
 -1.44915458e-01],
[ 2.54593326e-01,  4.98546558e-02,  4.92113010e-03,
  7.91883357e-02,  3.98064104e-01,  1.27615182e-01,
 -5.43413386e-02],
[ 3.98566589e-01, -1.31900722e-01,  1.04108639e-01,
  1.09696493e-01,  1.08133007e-01,  1.24932880e-01,
 -4.14726698e-02],
[ 3.80809842e-01, -1.07279362e-02,  7.88853232e-02,
  2.03762797e-01, -6.05040012e-02,  3.42438639e-01,
  7.66685412e-02],
[ 3.19815698e-01, -1.75654868e-02,  3.51328593e-01,
  3.15037549e-02,  2.83930766e-01,  9.90283697e-02,
 -3.02766690e-01],
[ 3.31405417e-01, -6.04905347e-02,  3.66609609e-01,
  3.78407093e-02,  6.08976868e-02,  2.97159598e-01,
 -4.32483058e-02],
[ 3.18934630e-01, -7.98822554e-02,  4.34991026e-01,
  1.22180337e-01,  5.80816309e-02,  4.37382987e-01,
 -1.10054358e-03],
[ 5.25460637e-01, -1.53108502e-01,  2.44469381e-01,
  1.37209535e-02,  4.15927944e-02,  2.94345918e-01,
  5.76168777e-02],
[ 6.05210594e-01, -6.19612023e-02,  3.12038776e-01,
  1.06680750e-01,  3.38630491e-01, -9.58537559e-02,
 -1.70190794e-01],
[ 5.74584842e-01, -1.44626773e-01,  4.17016149e-01,
  1.84513513e-01, -6.37025369e-02,  4.38148889e-01,
 -3.51065904e-02],
```

```
[ 5.16109735e-01, -8.84150983e-03,  3.28648288e-01,
  2.64141006e-01,  1.14098002e-01,  5.28520957e-01,
 -1.29692500e-01],
[ 6.14662482e-01, -4.08856089e-02, -2.29887711e-02,
  2.28795668e-01, -1.27627350e-01,  5.05755240e-01,
 -1.47295210e-01],
[ 3.22198327e-01, -2.38422190e-01,  3.79638302e-01,
  1.21660826e-01,  6.49225280e-02,  4.32463083e-01,
  1.80340949e-01],
[ 5.72025783e-01,  8.69978688e-02,  2.51883430e-01,
  7.14653897e-02,  8.99927841e-02,  5.77354049e-01,
 -3.20381459e-02],
[ 4.72358294e-01,  9.68056539e-02,  2.89294153e-01,
  1.06948832e-01,  3.46533381e-01,  2.75810599e-01,
  1.95122531e-01],
[ 4.90505633e-01,  8.82714088e-02,  2.71131937e-01,
  9.88179598e-02, -5.51994825e-02,  6.04947773e-01,
 -8.82628463e-02],
[ 5.41580610e-01, -3.45183571e-02,  5.14362489e-01,
  1.41658173e-01,  4.51961774e-01,  1.77539371e-02,
 -3.04371856e-01],
[ 5.19923718e-01,  5.32762994e-03,  1.53839390e-01,
 -1.07093126e-02,  4.82006834e-01,  8.45696461e-02,
 -1.80457010e-01],
[ 2.72978918e-01,  3.64536325e-02,  1.97745834e-01,
  3.02281503e-01,  2.46084714e-01,  3.59296454e-01,
 -2.83286704e-03],
[ 1.30894538e-01, -7.06155852e-02,  5.66196443e-02,
  9.10835684e-02,  1.47055337e-01,  4.42932317e-01,
 -1.52091857e-01],
[ 2.91733990e-01, -2.42926215e-02,  2.63095669e-01,
 -6.77047005e-02,  5.38140560e-01,  1.61957646e-02,
 -2.67216455e-01],
[ 1.47304528e-01, -2.76631943e-01,  1.42498001e-01,
  1.09063119e-01,  4.17546499e-01,  2.16980144e-01,
  2.30363630e-01],
[ 3.30089114e-01, -7.82535669e-02,  1.75721018e-01,
  1.22612179e-02,  5.07064332e-01,  7.66640071e-02,
 -3.73258471e-02],
[-4.39086634e-02, -3.20255151e-02,  7.18210454e-02,
  5.61271535e-02,  6.12729459e-01,  3.58971767e-01,
  7.40675547e-02],
[-3.03382897e-02, -1.49218268e-02, -1.00580206e-01,
  9.46223364e-02,  3.34203549e-01,  6.10204523e-01,
  1.89532342e-02],
[-2.61965358e-01, -1.47595588e-01,  1.27063492e-01,
  3.85796357e-01,  1.45104406e+00, -5.93058604e-01,
  1.19375968e-01],
```

```
[-2.83519914e-01,  1.07428579e-02, -3.83449199e-01,
  5.25149255e-02,  6.89603942e-01,  9.35356800e-02,
  4.58760348e-02],
[-2.91480479e-01, -1.31055332e-01,  5.09766174e-01,
  2.98725699e-01,  1.08246174e+00, -3.03449964e-01,
 -1.24062888e-01],
[-3.05672835e-01,  1.00569937e-01, -2.71190369e-01,
  1.80270113e-01,  1.19565296e+00, -5.26343647e-01,
 -5.96822567e-02],
[-3.38482746e-01, -2.53891405e-01, -2.40350064e-02,
  1.43059838e-01,  6.18609612e-01,  8.15421308e-03,
 -2.28562159e-01],
[-2.57183461e-01, -1.94401198e-01, -1.70357035e-01,
  2.62268122e-01,  8.24013911e-01, -2.01780682e-01,
  5.74041823e-02],
[-3.32829289e-01, -3.41988751e-01, -5.19077490e-01,
  2.41961812e-01,  4.67997230e-01,  1.07741122e-01,
 -7.64982219e-02],
[-6.18551217e-01, -1.67361395e-01,  1.33822165e-02,
  2.92508977e-02,  6.94249490e-01, -2.58415301e-02,
  2.82503705e-02],
[-4.37327947e-01,  6.50926191e-02, -8.31249911e-01,
  3.38338934e-02,  8.67506591e-01, -3.23565981e-01,
 -1.47269037e-01],
[-2.25150730e-01,  3.98077690e-01, -1.33389849e-01,
 -1.02695310e-02,  5.37371632e-01, -4.81168465e-01,
 -1.03621388e-01],
[-8.72136008e-01, -1.58784435e-01, -5.54419431e-01,
  1.66665098e-01,  2.33747636e-01,  2.71067900e-01,
 -4.80252684e-01],
[-8.75789843e-01, -1.21847706e-01, -8.57359711e-02,
  1.95995502e-01,  6.78238667e-01,  1.76614610e-01,
 -6.43350487e-01],
[-8.28113498e-02, -4.91738168e-01, -3.63169144e-01,
 -1.03120868e-02,  4.20089775e-02, -1.33087002e-01,
 -3.93192504e-01],
[-6.32622048e-01, -4.39375337e-02, -8.66230541e-01,
 -2.86900421e-01, -4.63936510e-01,  5.45660012e-01,
 -4.32556113e-01],
[-9.06321579e-01, -6.24305979e-01, -2.31126502e-01,
 -1.93309365e-01,  3.93069398e-01, -4.22415369e-01,
 -3.43844191e-01],
[-6.39492211e-01, -2.32577721e-01, -5.67878119e-01,
 -2.73155169e-01, -2.03648931e-01,  4.91285104e-01,
 -7.34128420e-01],
[-3.07724118e-01, -5.42936987e-01, -7.49275198e-01,
 -3.96754650e-01,  2.67645714e-01, -3.44925492e-01,
 -1.10834721e+00],
```

```
        [ 3.07464489e-01,  1.79594977e-02, -1.35975139e-01,
         -6.40330339e-01, -4.86138145e-01,  4.23215531e-01,
         -1.03719523e+00],
        [-1.02573137e+00, -2.97589391e-01, -7.05897268e-01,
         -7.63634443e-01, -1.08479291e-01,  7.62764009e-02,
         -7.97500381e-01],
        [-3.85227385e-01, -2.30017452e-01, -6.45425993e-01,
         -1.16284508e-01,  4.06333820e-02, -5.81843773e-01,
         -1.53524659e+00],
        [-1.13830241e+00, -5.57554409e-01, -8.54072008e-01,
         -1.21086477e+00, -3.02430700e-01, -6.99851949e-01,
         -1.26996572e+00],
        [ 1.17704374e-01, -6.26100496e-01, -6.88172554e-01,
         -1.18763191e+00, -6.02111477e-01, -8.69059566e-01,
         -1.56874365e+00],
        [-6.47237662e-01, -8.57311843e-01,  7.37059549e-02,
         -5.59646437e-01, -2.32711375e-01, -5.97680069e-01,
         -1.27567092e+00],
        [-6.65344348e-01, -9.37661571e-01, -5.26356965e-01,
         -9.69645041e-01, -5.87003467e-01, -8.45579057e-01,
         -2.20839244e+00],
        [-4.97598802e-01, -8.22280769e-01, -4.48817079e-01,
         -8.26538139e-01, -8.57610655e-01, -1.04098300e+00,
         -2.26377329e+00],
        [-5.08168608e-01, -8.06922527e-01, -4.36342053e-01,
         -1.27960362e+00, -8.82122830e-01, -1.06256379e+00,
         -2.24423439e+00],
        [-7.81018162e-01, -4.61679786e-01, -5.52422649e-01,
         -9.45510607e-01, -3.31922952e-01, -6.08311739e-01,
         -1.92028723e+00],
        [-3.52249604e-01,  2.10697280e-01, -1.62212670e-01,
         -3.14733259e-01,  3.16274756e-01, -5.36097799e-01,
         -8.80876383e-01]])

In [90]: model.predict(X)

Out[90]: array([20, 20, 20, ..., 45, 37, 37], dtype=int64)

In [87]: X

Out[87]:      salary_num  edu_num  workclass_num  sex_num  rel_num  ms_num  nc_num
        0           0        0              0        1        0       0       1
        1           0        0              0        1        0       0       1
        2           0        0              0        1        0       0       1
        3           0        0              0        1        0       0       1
        4           0        0              0        1        1       1       1
        5           0        0              0        1        0       0       1
        6           0        0              1        0        1       1       1
        7           0        0              0        0        1       1       1
```

8	0	0	0	0	1	1	1
9	0	0	0	0	1	1	1
10	0	0	0	0	1	1	1
11	0	0	0	0	1	1	1
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	1
15	0	0	0	0	0	0	0
16	0	0	0	0	0	0	1
17	0	0	0	0	1	1	1
18	0	0	0	0	0	0	1
19	0	0	0	0	0	1	1
20	0	0	0	0	0	0	1
21	0	0	0	0	0	0	1
22	0	0	0	0	1	1	0
23	0	0	0	0	0	0	1
24	0	0	0	0	0	0	1
25	0	0	0	0	0	0	1
26	0	0	0	0	0	0	1
27	0	0	0	1	0	0	0
28	0	0	0	1	0	0	1
29	0	0	0	1	0	0	1
...	...	...	...	...	...	...	...
32531	1	1	0	0	1	1	1
32532	1	1	0	0	1	1	1
32533	1	1	0	0	1	1	1
32534	1	1	0	0	1	1	1
32535	1	1	0	0	1	1	1
32536	1	1	0	0	1	1	1
32537	1	1	0	0	1	1	1
32538	1	1	0	0	1	1	1
32539	1	1	0	0	1	1	1
32540	1	1	0	0	1	1	1
32541	1	1	0	0	1	1	1
32542	1	1	0	0	1	1	1
32543	1	1	0	0	1	1	1
32544	1	1	0	0	1	1	1
32545	1	1	0	0	1	1	1
32546	1	1	0	0	1	1	1
32547	1	1	0	0	1	1	1
32548	1	1	0	0	1	1	1
32549	1	1	0	0	1	1	1
32550	1	1	0	0	1	1	1
32551	1	1	0	0	1	1	1
32552	1	1	1	0	0	0	1
32553	1	1	0	0	0	0	1
32554	1	1	0	0	0	0	1
32555	1	1	0	0	1	1	1

```
       32556           1           1              0           0           0           0           1
       32557           1           1              0           0           0           0           1
       32558           1           1              1           0           0           0           1
       32559           1           1              0           0           0           0           1
       32560           1           1              0           0           0           0           1

       [32561 rows x 7 columns]

In [91]: accuracy_score(adult.age, adult.pred.round())

Out[91]: 0.044347532323945824

In [92]: # fit a linear regression model and store the predictions
         feature_cols = ['salary_num', 'edu_num', 'workclass_num', 'sex_num', 'ms_num']
         X = adult[feature_cols]
         y = adult.age

         model = Model()
         model.fit(X, y)
         adult['pred'] = model.predict(X)


         from sklearn.metrics import accuracy_score, confusion_matrix
         accuracy_score(adult.age, adult.pred.round())
         confusion_matrix(adult.age, adult.pred.round())

Out[92]: array([[  0,    0,  77, ...,    0,    0,    0],
                [  0,    0, 107, ...,    0,    0,    0],
                [  0,    0, 133, ...,    0,    0,    0],
                ...,
                [  0,    0,    0, ...,    0,    0,    0],
                [  0,    0,    1, ...,    0,    0,    0],
                [  0,    0,    4, ...,    0,    0,    0]], dtype=int64)

In [93]: from sklearn.metrics import classification_report
         print(classification_report(adult.age, adult.pred.round()))

              precision    recall  f1-score   support

          17       0.00      0.00      0.00       395
          18       0.00      0.00      0.00       550
          19       0.07      0.19      0.10       712
          20       0.04      0.31      0.08       753
          21       0.00      0.00      0.00       720
          22       0.00      0.00      0.00       765
          23       0.05      0.52      0.09       877
          24       0.00      0.00      0.00       798
          25       0.00      0.00      0.00       841
          26       0.00      0.00      0.00       785
```

70

27	0.00	0.00	0.00	835
28	0.00	0.00	0.00	867
29	0.00	0.00	0.00	813
30	0.00	0.00	0.00	861
31	0.04	0.07	0.05	888
32	0.00	0.00	0.00	828
33	0.00	0.00	0.00	875
34	0.00	0.00	0.00	886
35	0.04	0.21	0.06	876
36	0.03	0.02	0.03	898
37	0.04	0.16	0.06	858
38	0.04	0.07	0.05	827
39	0.04	0.03	0.04	816
40	0.01	0.00	0.00	794
41	0.02	0.00	0.01	808
42	0.00	0.00	0.00	780
43	0.04	0.04	0.04	770
44	0.00	0.00	0.00	724
45	0.06	0.01	0.02	734
46	0.06	0.07	0.06	737
47	0.05	0.06	0.05	708
48	0.00	0.00	0.00	543
49	0.00	0.00	0.00	577
50	0.00	0.00	0.00	602
51	0.00	0.00	0.00	595
52	0.00	0.00	0.00	478
53	0.00	0.00	0.00	464
54	0.00	0.00	0.00	415
55	0.00	0.00	0.00	419
56	0.00	0.00	0.00	366
57	0.00	0.00	0.00	358
58	0.00	0.00	0.00	366
59	0.00	0.00	0.00	355
60	0.00	0.00	0.00	312
61	0.00	0.00	0.00	300
62	0.00	0.00	0.00	258
63	0.00	0.00	0.00	230
64	0.00	0.00	0.00	208
65	0.00	0.00	0.00	178
66	0.00	0.00	0.00	150
67	0.00	0.00	0.00	151
68	0.00	0.00	0.00	120
69	0.00	0.00	0.00	108
70	0.00	0.00	0.00	89
71	0.00	0.00	0.00	72
72	0.00	0.00	0.00	67
73	0.00	0.00	0.00	64
74	0.00	0.00	0.00	51

```
         75      0.00      0.00      0.00        45
         76      0.00      0.00      0.00        46
         77      0.00      0.00      0.00        29
         78      0.00      0.00      0.00        23
         79      0.00      0.00      0.00        22
         80      0.00      0.00      0.00        22
         81      0.00      0.00      0.00        20
         82      0.00      0.00      0.00        12
         83      0.00      0.00      0.00         6
         84      0.00      0.00      0.00        10
         85      0.00      0.00      0.00         3
         86      0.00      0.00      0.00         1
         87      0.00      0.00      0.00         1
         88      0.00      0.00      0.00         3
         90      0.00      0.00      0.00        43

avg / total      0.01      0.04      0.02     32561


C:\Users\Erin\Anaconda3\lib\site-packages\sklearn\metrics\classification.py:1135: UndefinedMetr
  'precision', 'predicted', average, warn_for)


In [94]: accuracy_score(adult.age, adult.pred.round())

Out[94]: 0.04471607137372931

In [95]: # fit a linear regression model and store the predictions
         feature_cols = ['age', 'edu_num', 'workclass_num', 'sex_num', 'ms_num']
         X = adult[feature_cols]
         y = adult.salary_num

         model = Model()
         model.fit(X, y)
         adult['pred'] = model.predict(X)


         from sklearn.metrics import accuracy_score, confusion_matrix
         accuracy_score(adult.salary_num, adult.pred.round())
         confusion_matrix(adult.salary_num, adult.pred.round())

Out[95]: array([[22819,  1901],
                [ 6191,  1650]], dtype=int64)

In [96]: from sklearn.metrics import classification_report
         print(classification_report(adult.salary_num, adult.pred.round()))

             precision    recall  f1-score   support
```

```
            0       0.79      0.92      0.85      24720
            1       0.46      0.21      0.29       7841

avg / total         0.71      0.75      0.71      32561
```

In [97]: accuracy_score(adult.salary_num, adult.pred.round())

Out[97]: 0.7514818340960044