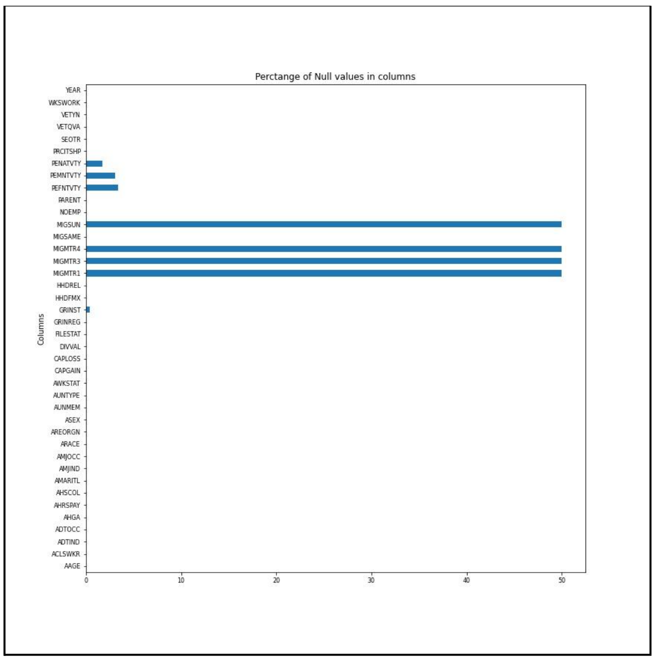
Machine Learning : CSE343

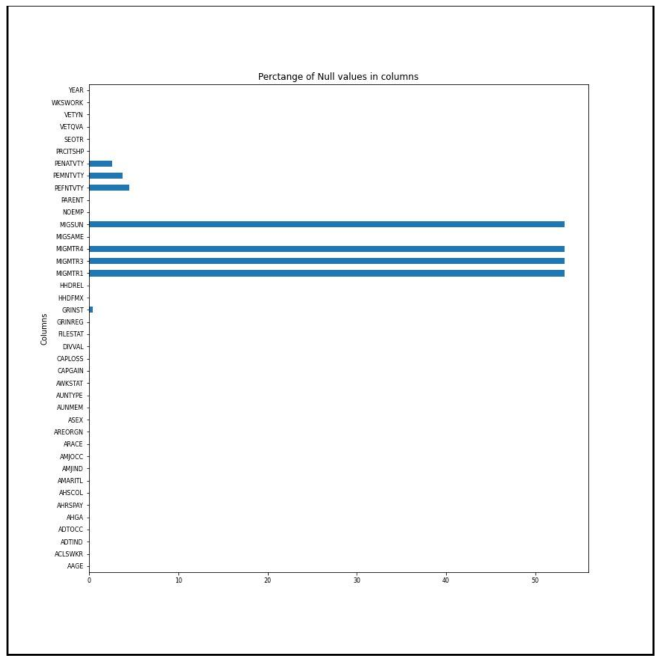
Assignment - 3

- Abhimanyu Gupta
(2019226)

Plots:

Percentage of Nulls

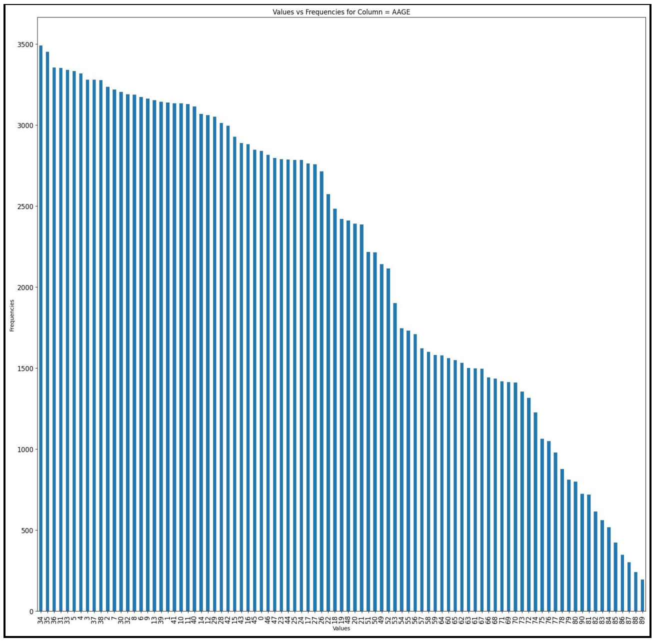
<table>
<tr><th>population.csv</th><th>more_than_50K.csv</th></tr>
<tr><td>


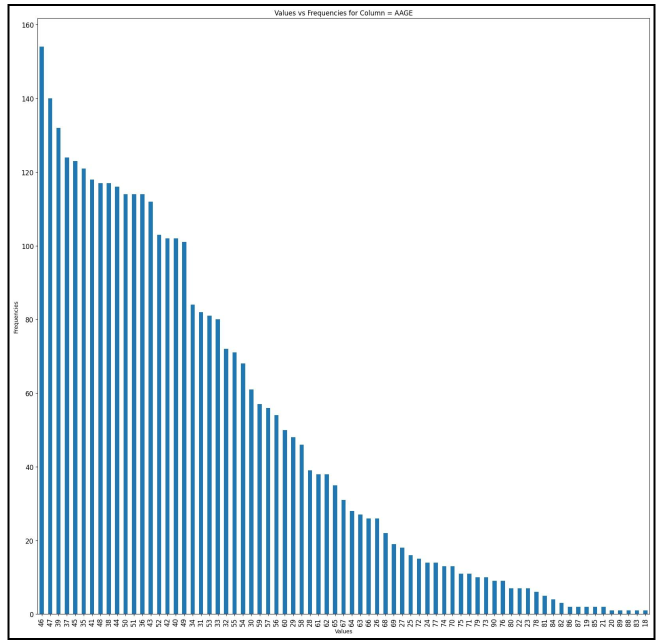
</td><td>


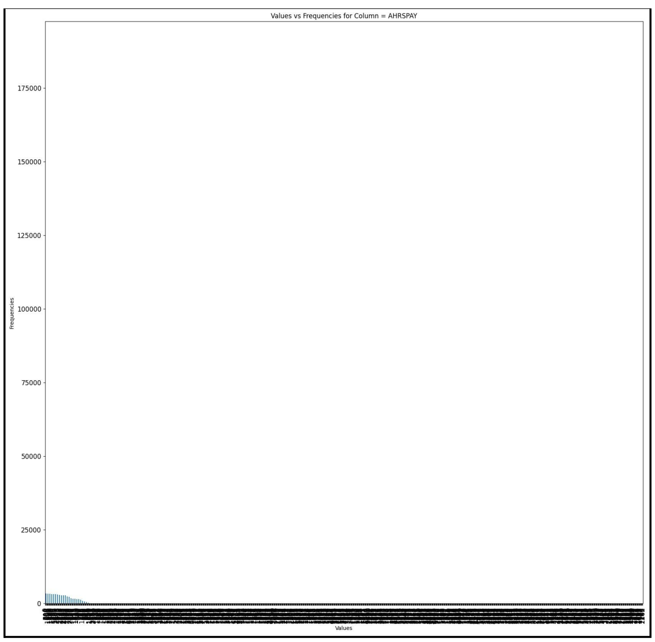
</td></tr>
</table>

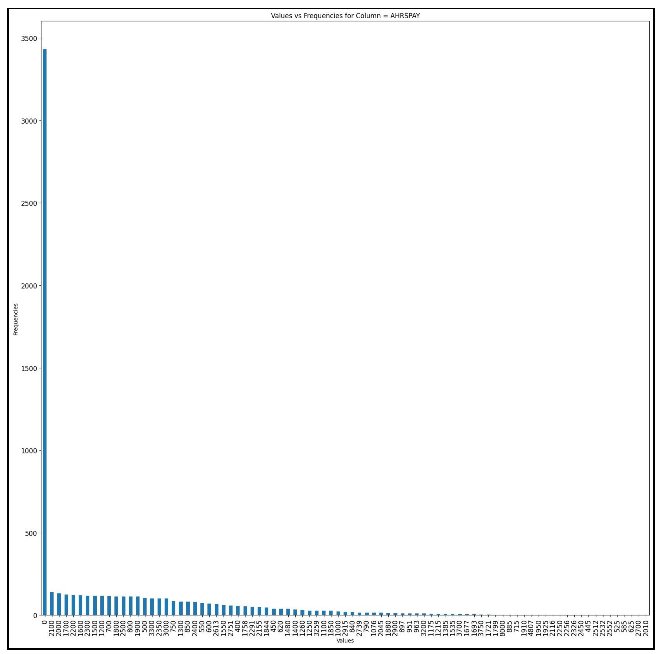
Numerical Data

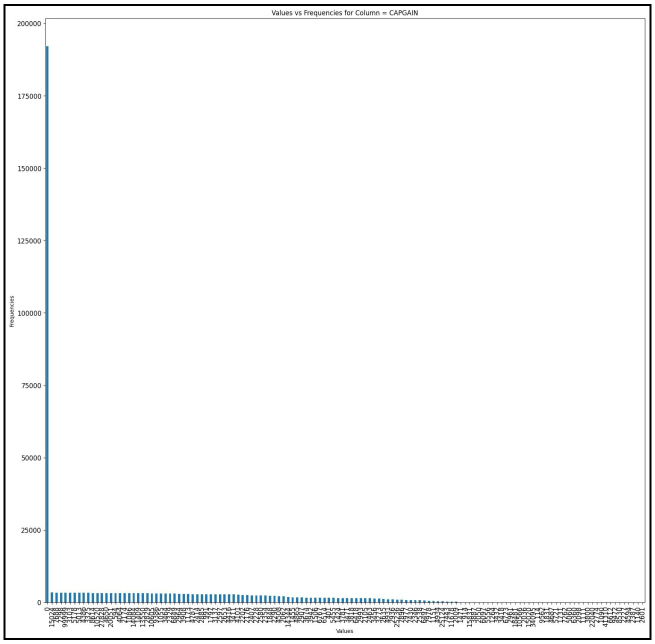
population.csv	more_than_50K.csv

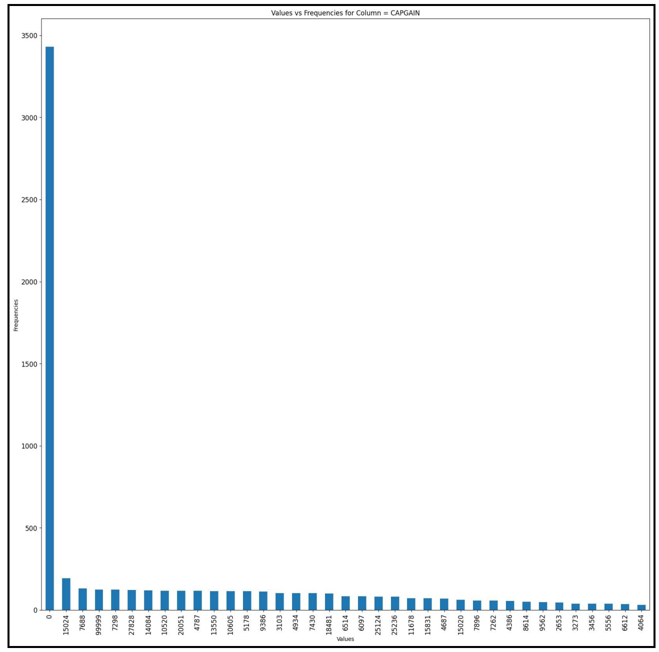

Values vs Frequencies for Column = AAGE
Frequencies
Values

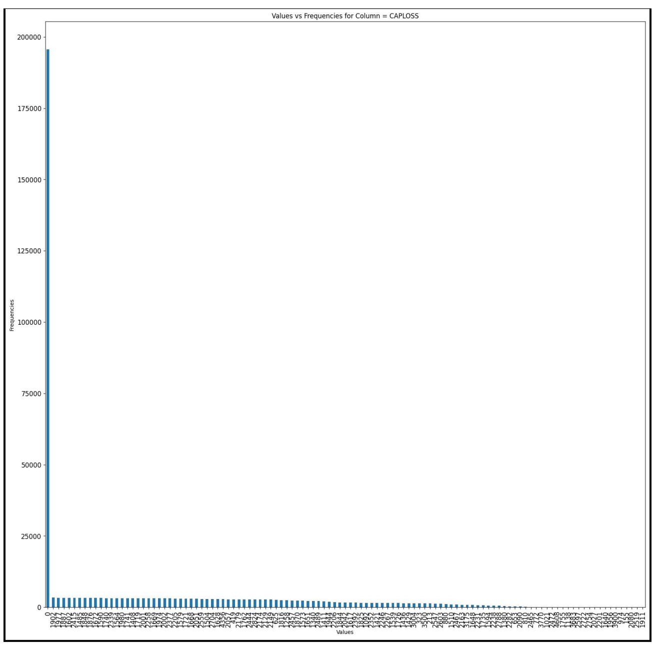

Values vs Frequencies for Column = AAGE
Frequencies
Values

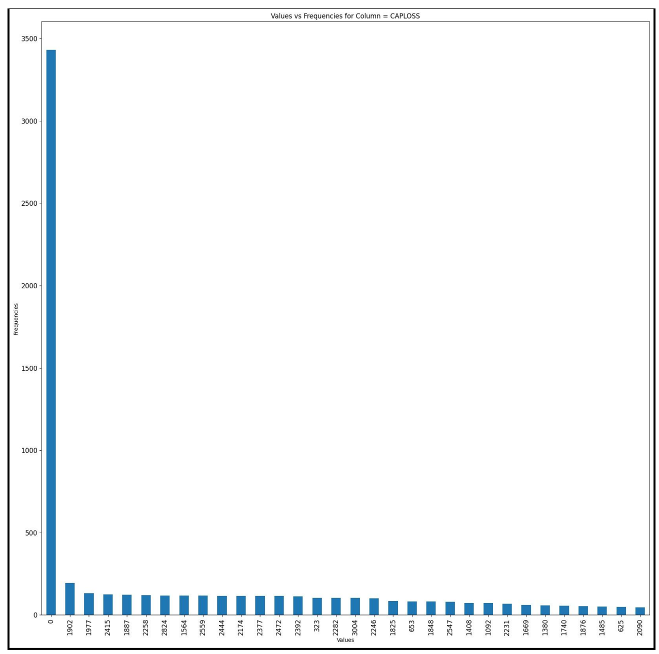

Values vs Frequencies for Column = AHRSPAY
Frequencies
Values

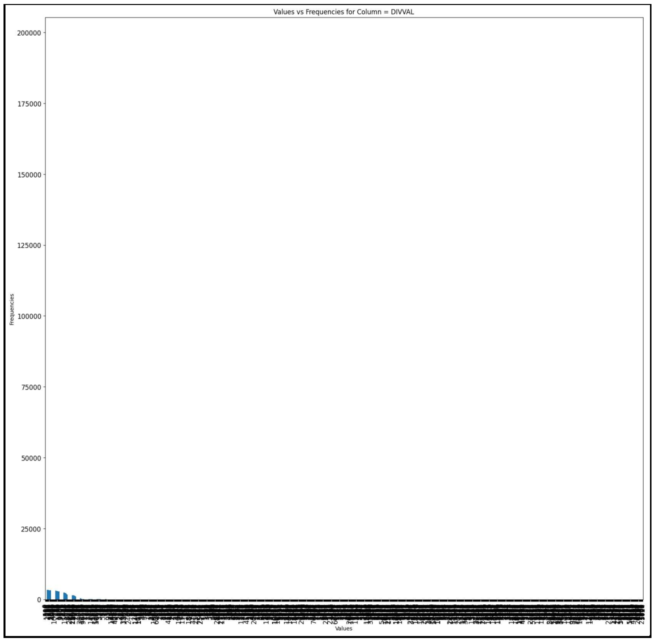

Values vs Frequencies for Column = AHRSPAY
Frequencies
Values

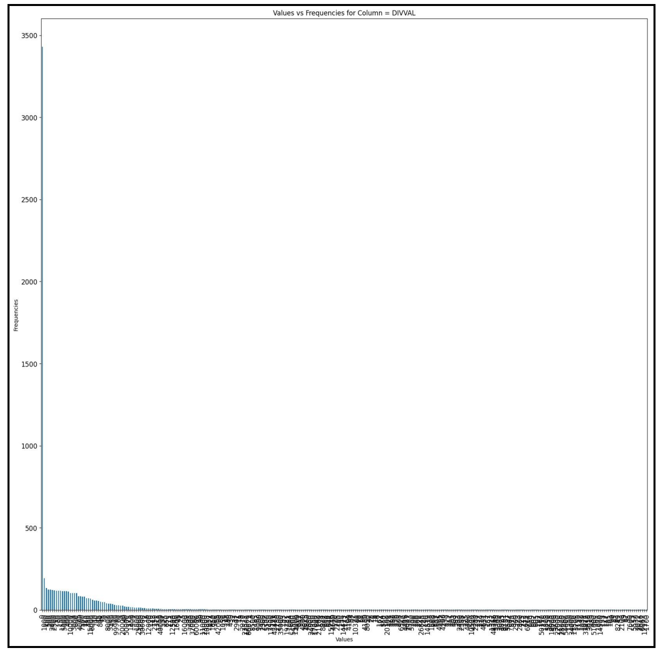

Values vs Frequencies for Column = CAPGAIN
Frequencies
Values
200000
175000
150000
125000
100000
75000
50000
25000
0

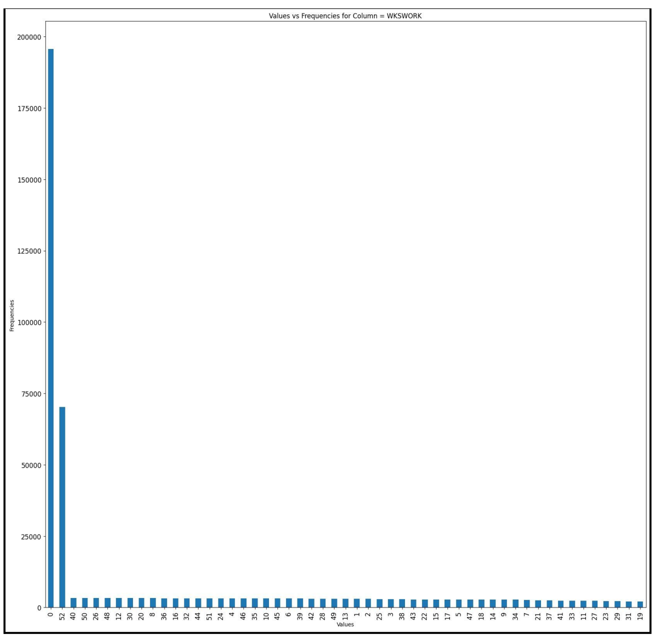

Values vs Frequencies for Column = CAPGAIN
Frequencies
Values
3500
3000
2500
2000
1500
1000
500
0
0
15024
7688
99999
7298
27828
14084
10520
20051
4787
13550
10605
5178
9386
3103
4934
7430
18481
6514
6097
25124
25236
11678
15831
4687
15020
7896
7262
4386
8614
9562
2653
3273
3456
5556
6612
4064

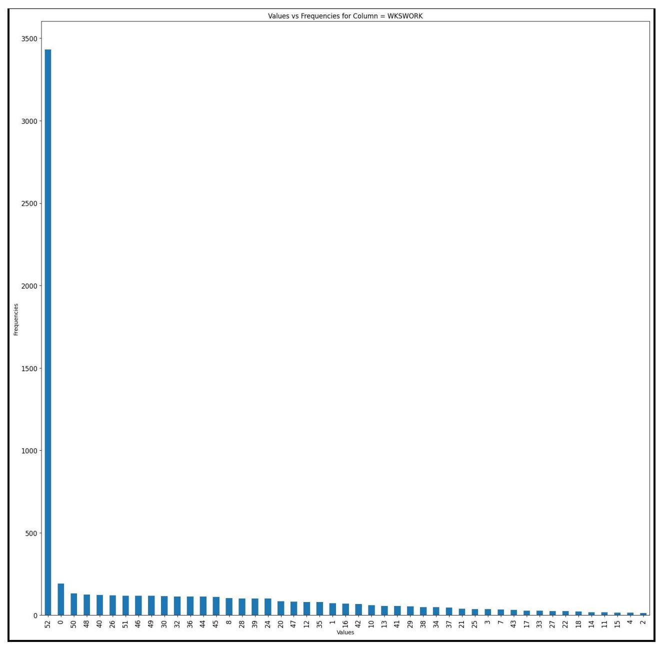

Values vs Frequencies for Column = CAPLOSS
Frequencies
Values
200000
175000
150000
125000
100000
75000
50000
25000
0


Values vs Frequencies for Column = CAPLOSS
Frequencies
Values
3500
3000
2500
2000
1500
1000
500
0
0
1902
1977
2415
1887
2258
2824
1564
2559
2444
2174
2377
2472
2392
323
2282
3004
2246
1825
653
1848
2547
1408
1092
2231
1669
1380
1740
1876
1485
625
2090



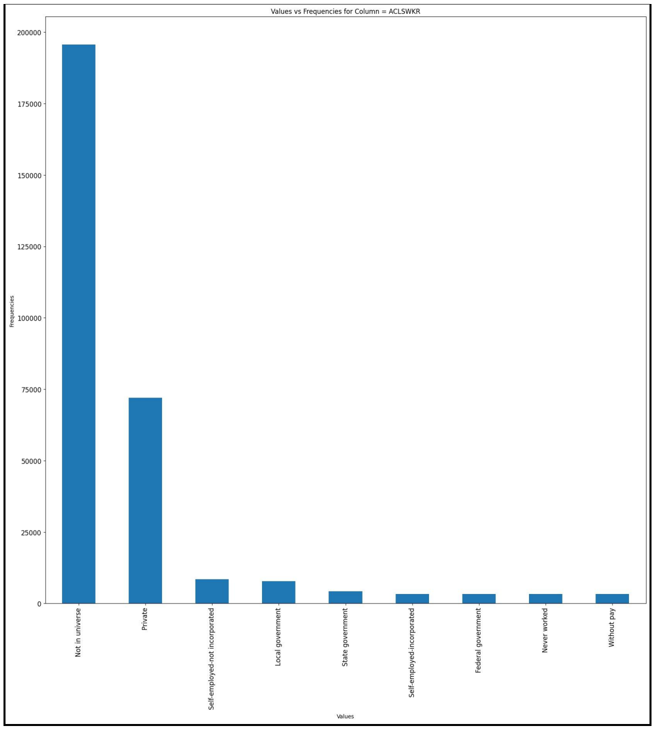


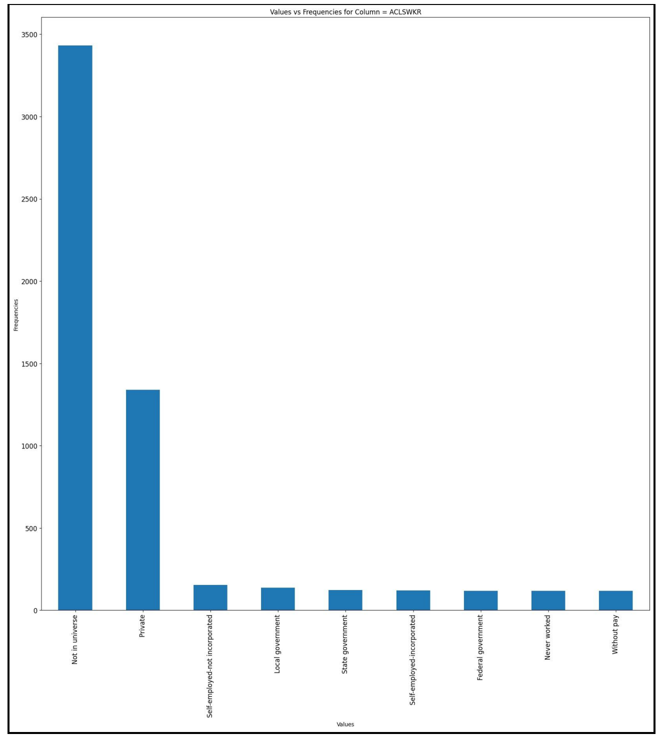


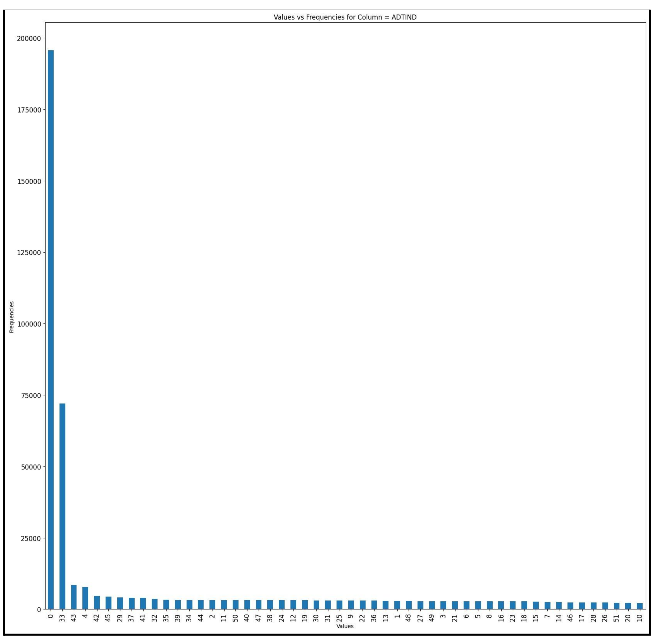


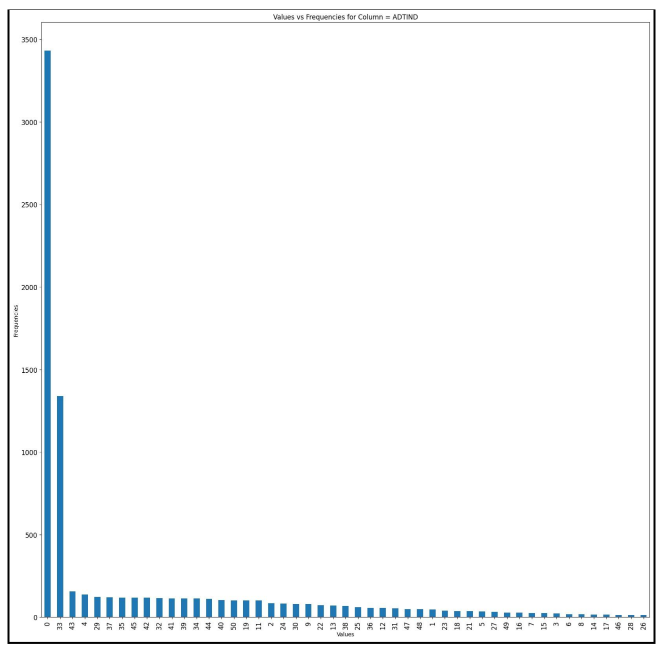
Categorical Data

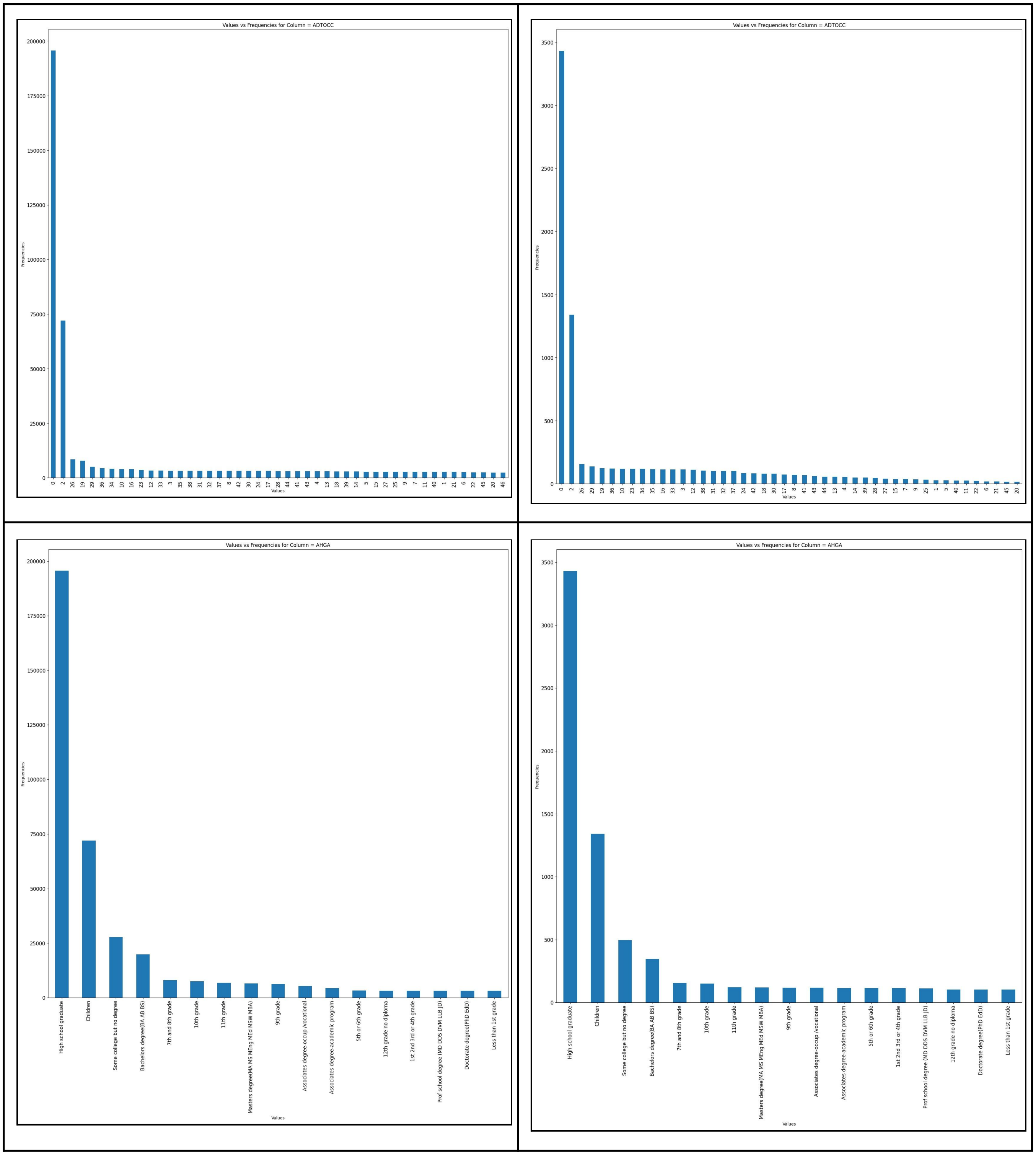
population.csv	**more_than_50K.csv**

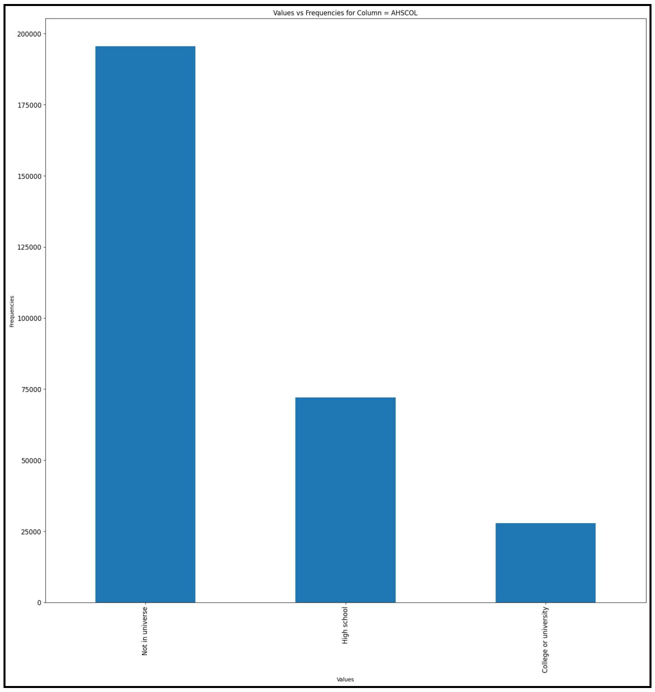

Values vs Frequencies for Column = ACLSWKR
Frequencies
Values
0
25000
50000
75000
100000
125000
150000
175000
200000
Not in universe
Private
Self-employed-not incorporated
Local government
State government
Self-employed-incorporated
Federal government
Never worked
Without pay

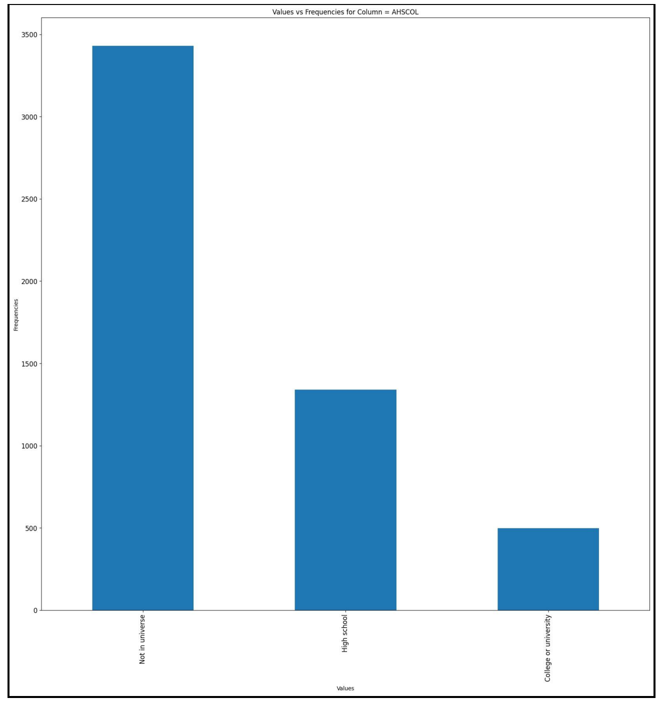

Values vs Frequencies for Column = ACLSWKR
Frequencies
Values
0
500
1000
1500
2000
2500
3000
3500
Not in universe
Private
Self-employed-not incorporated
Local government
State government
Self-employed-incorporated
Federal government
Never worked
Without pay

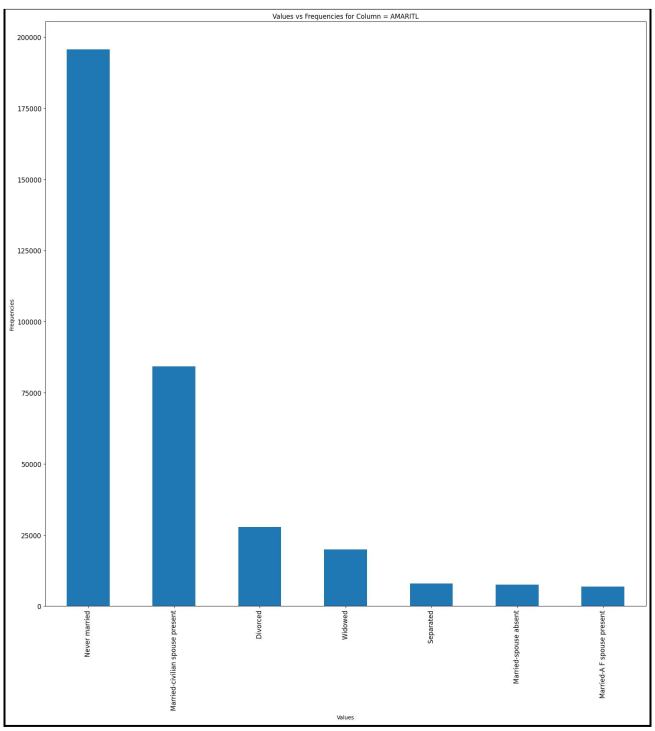

Values vs Frequencies for Column = ADTIND
Frequencies
Values
0
25000
50000
75000
100000
125000
150000
175000
200000

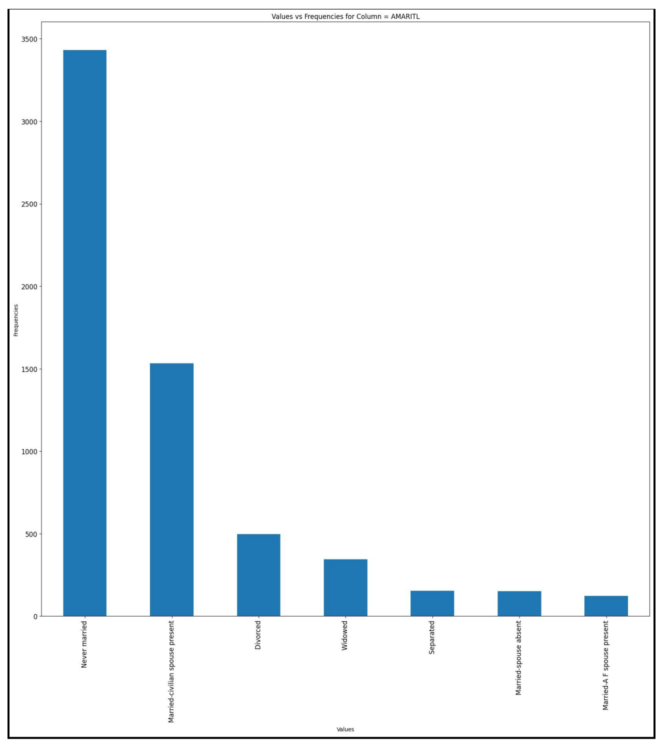

Values vs Frequencies for Column = ADTIND
Frequencies
Values
0
500
1000
1500
2000
2500
3000
3500

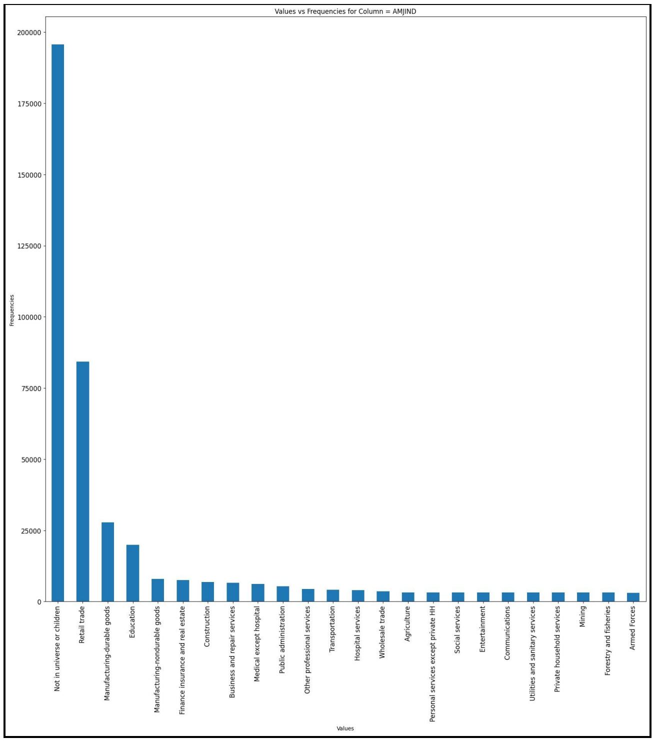

Values vs Frequencies for Column = ADTOCC
Frequencies
Values
Values vs Frequencies for Column = ADTOCC
Frequencies
Values
Values vs Frequencies for Column = AHGA
Frequencies
High school graduate
Children
Some college but no degree
Bachelors degree(BA AB BS)
7th and 8th grade
10th grade
11th grade
Masters degree(MA MS MEng MEd MSW MBA)
9th grade
Associates degree-occup /vocational
Associates degree-academic program
5th or 6th grade
12th grade no diploma
1st 2nd 3rd or 4th grade
Prof school degree (MD DDS DVM LLB JD)
Doctorate degree(PhD EdD)
Less than 1st grade
Values
Values vs Frequencies for Column = AHGA
Frequencies
Values

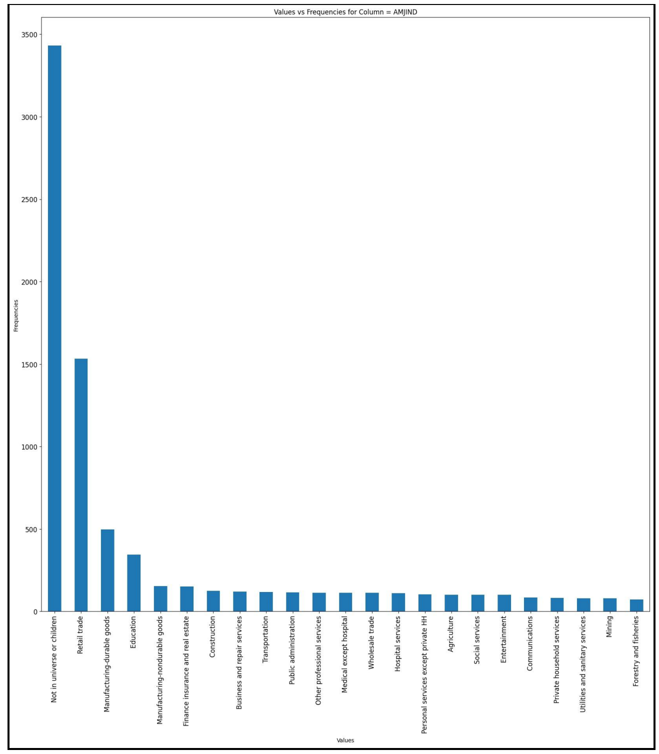

Values vs Frequencies for Column = AHSCOL
Frequencies
Values
Not in universe
High school
College or university

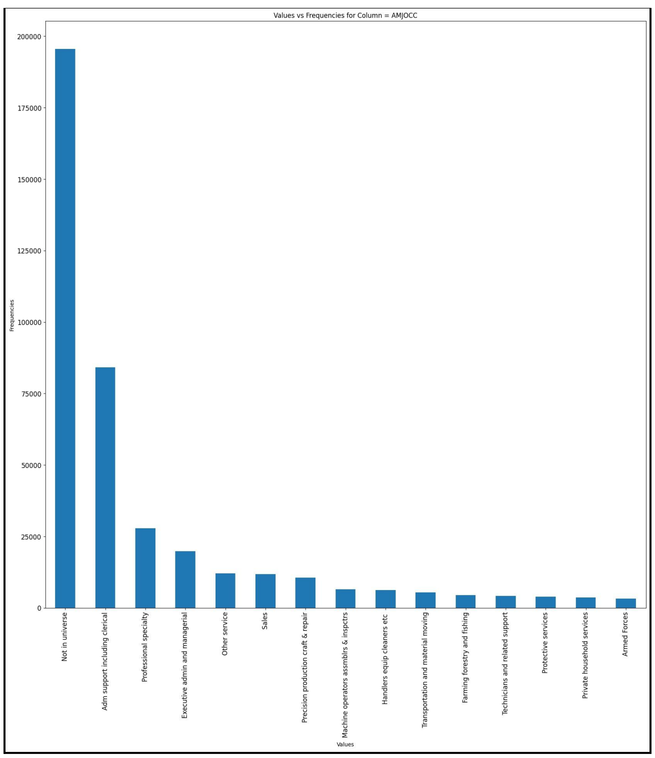

Values vs Frequencies for Column = AHSCOL
Frequencies
Values
Not in universe
High school
College or university

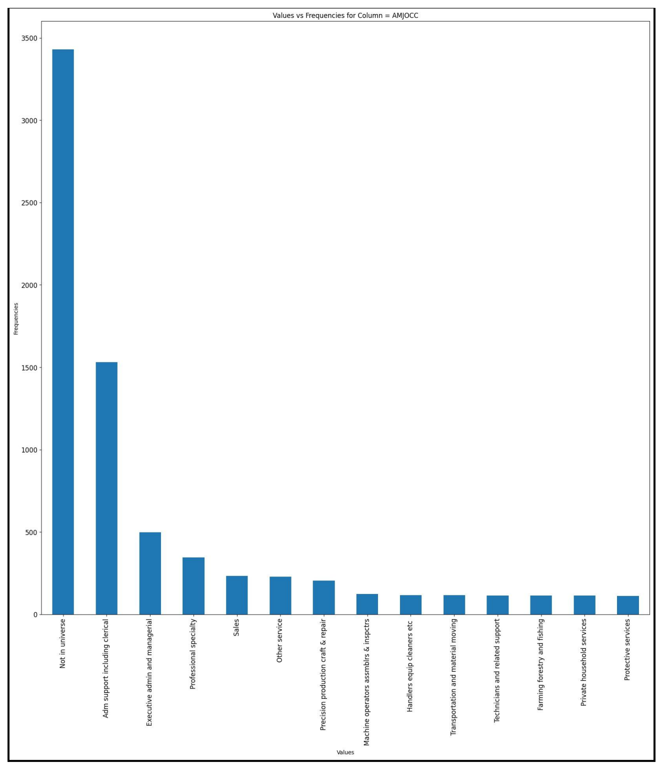

Values vs Frequencies for Column = AMARITL
Frequencies
Values
Never married
Married-civilian spouse present
Divorced
Widowed
Separated
Married-spouse absent
Married-A F spouse present

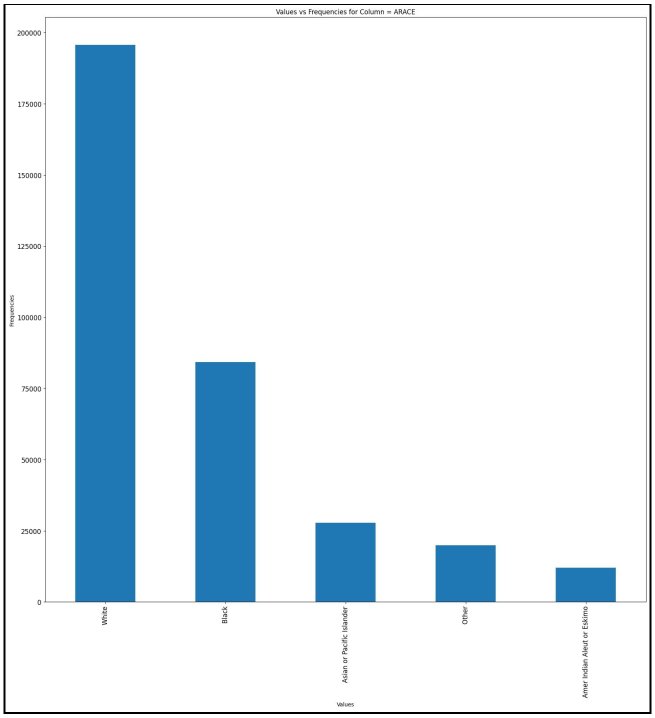

Values vs Frequencies for Column = AMARITL
Frequencies
Values
Never married
Married-civilian spouse present
Divorced
Widowed
Separated
Married-spouse absent
Married-A F spouse present

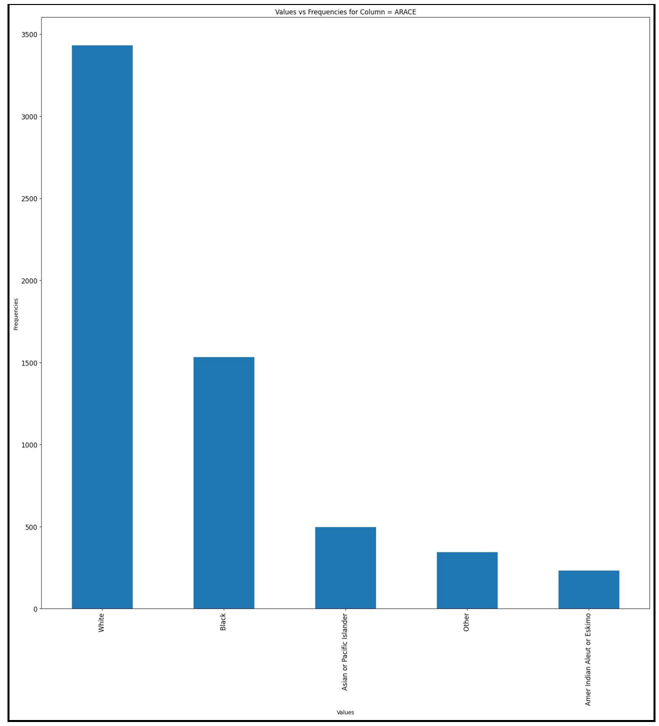

Values vs Frequencies for Column = AMJIND
0
25000
50000
75000
100000
125000
150000
175000
200000
Frequencies
Not in universe or children
Retail trade
Manufacturing-durable goods
Education
Manufacturing-nondurable goods
Finance insurance and real estate
Construction
Business and repair services
Medical except hospital
Public administration
Other professional services
Transportation
Hospital services
Wholesale trade
Agriculture
Personal services except private HH
Social services
Entertainment
Communications
Utilities and sanitary services
Private household services
Mining
Forestry and fisheries
Armed Forces
Values

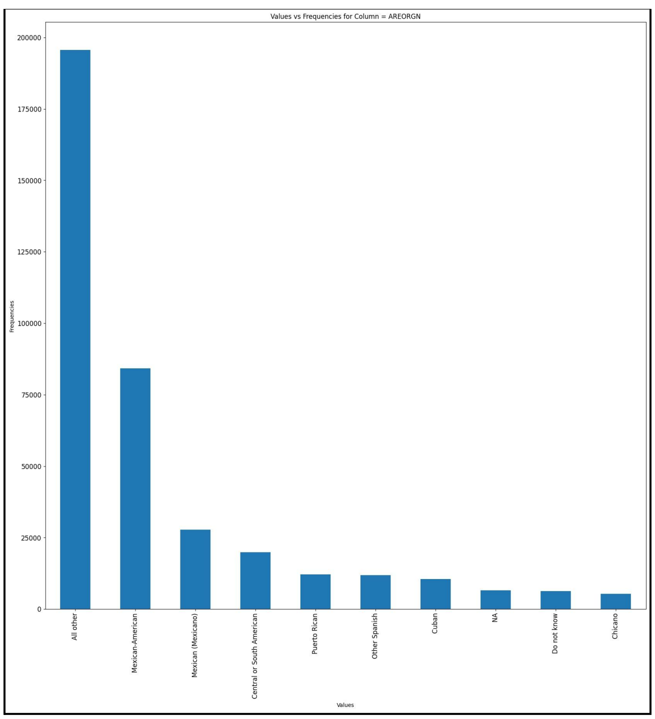

Values vs Frequencies for Column = AMJIND
0
500
1000
1500
2000
2500
3000
3500
Frequencies
Not in universe or children
Retail trade
Manufacturing-durable goods
Education
Manufacturing-nondurable goods
Finance insurance and real estate
Construction
Business and repair services
Transportation
Public administration
Other professional services
Medical except hospital
Wholesale trade
Hospital services
Personal services except private HH
Agriculture
Social services
Entertainment
Communications
Private household services
Utilities and sanitary services
Mining
Forestry and fisheries
Values

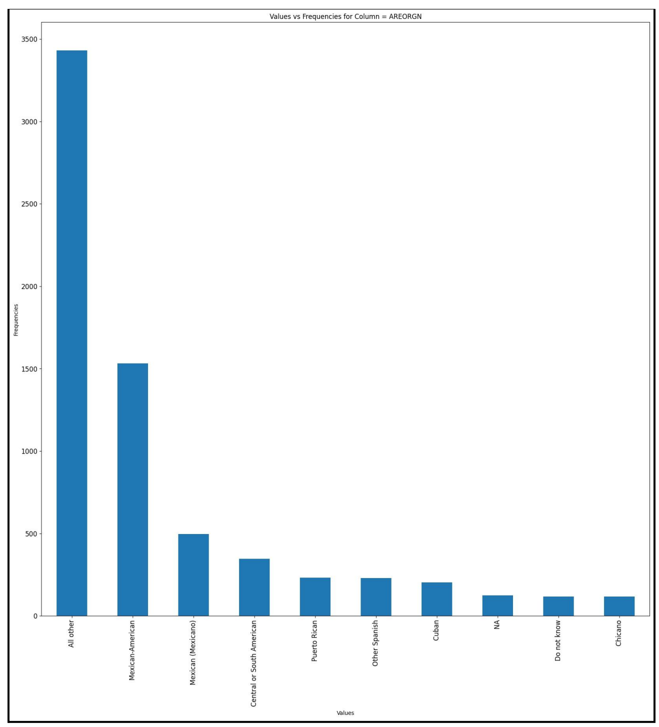

Values vs Frequencies for Column = AMJOCC
0
25000
50000
75000
100000
125000
150000
175000
200000
Frequencies
Not in universe
Adm support including clerical
Professional specialty
Executive admin and managerial
Other service
Sales
Precision production craft & repair
Machine operators assmblrs & inspctrs
Handlers equip cleaners etc
Transportation and material moving
Farming forestry and fishing
Technicians and related support
Protective services
Private household services
Armed Forces
Values

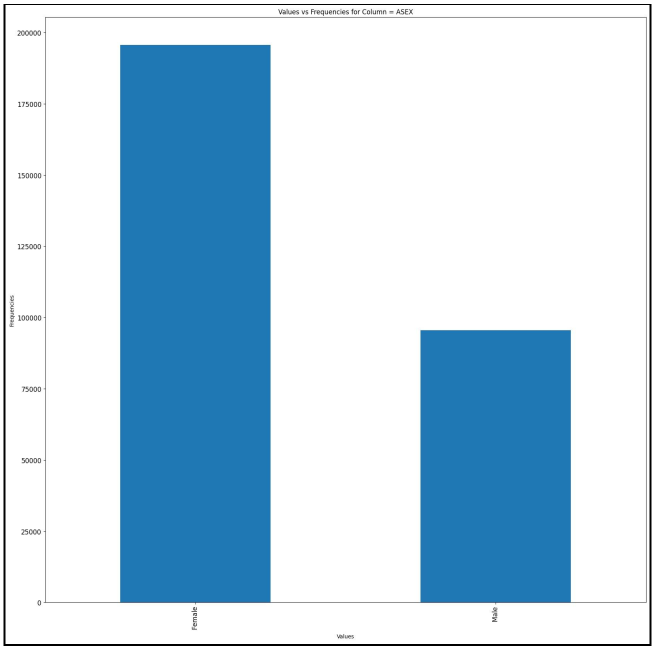

Values vs Frequencies for Column = AMJOCC
0
500
1000
1500
2000
2500
3000
3500
Frequencies
Not in universe
Adm support including clerical
Executive admin and managerial
Professional specialty
Sales
Other service
Precision production craft & repair
Machine operators assmblrs & inspctrs
Handlers equip cleaners etc
Transportation and material moving
Technicians and related support
Farming forestry and fishing
Private household services
Protective services
Values

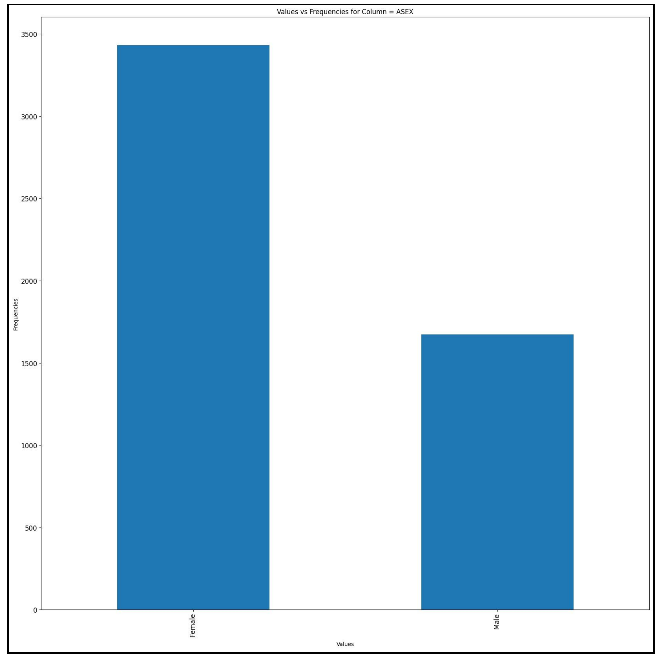

Values vs Frequencies for Column = ARACE
Frequencies
Values
0
25000
50000
75000
100000
125000
150000
175000
200000
White
Black
Asian or Pacific Islander
Other
Amer Indian Aleut or Eskimo

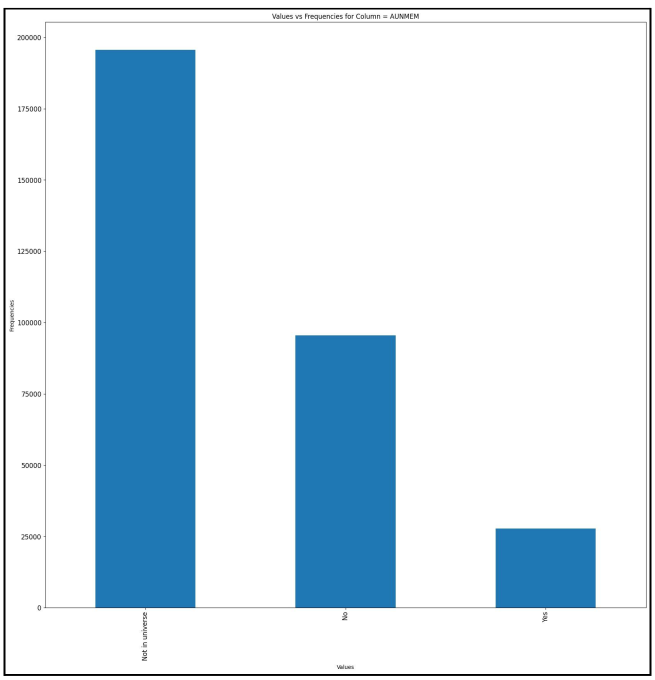

Values vs Frequencies for Column = ARACE
Frequencies
Values
0
500
1000
1500
2000
2500
3000
3500
White
Black
Asian or Pacific Islander
Other
Amer Indian Aleut or Eskimo

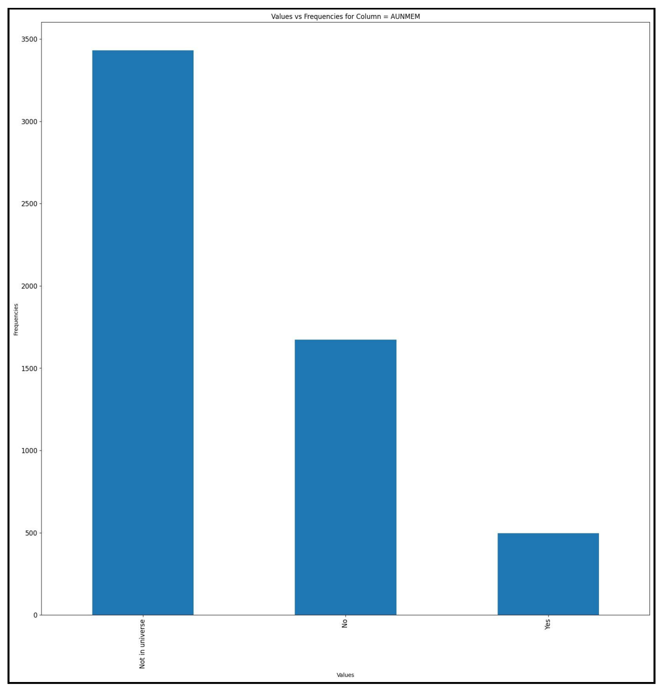

Values vs Frequencies for Column = AREORGN
Frequencies
Values
0
25000
50000
75000
100000
125000
150000
175000
200000
All other
Mexican-American
Mexican (Mexicano)
Central or South American
Puerto Rican
Other Spanish
Cuban
NA
Do not know
Chicano

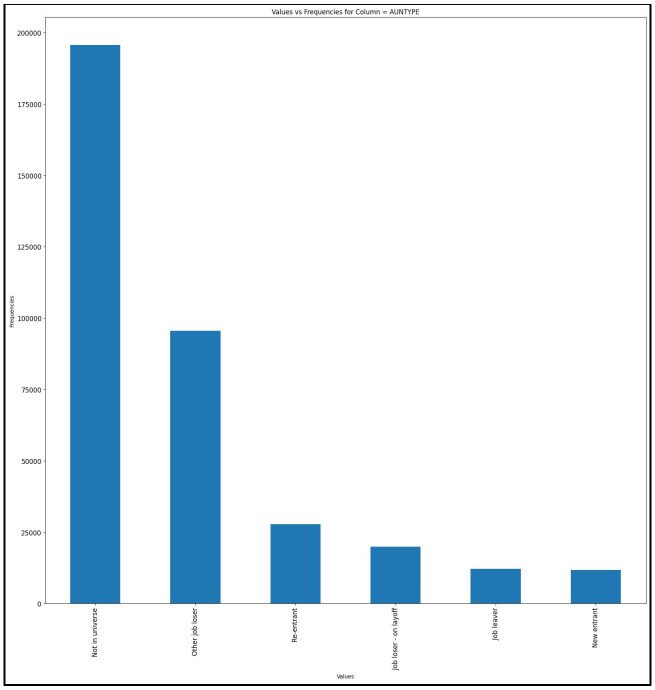

Values vs Frequencies for Column = AREORGN
Frequencies
Values
0
500
1000
1500
2000
2500
3000
3500
All other
Mexican-American
Mexican (Mexicano)
Central or South American
Puerto Rican
Other Spanish
Cuban
NA
Do not know
Chicano

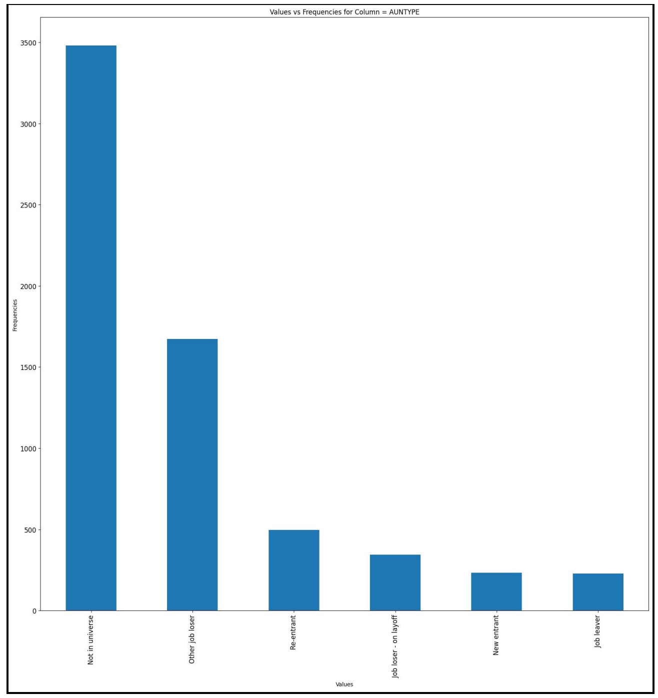

Values vs Frequencies for Column = ASEX
Frequencies
0
25000
50000
75000
100000
125000
150000
175000
200000
Female
Male
Values

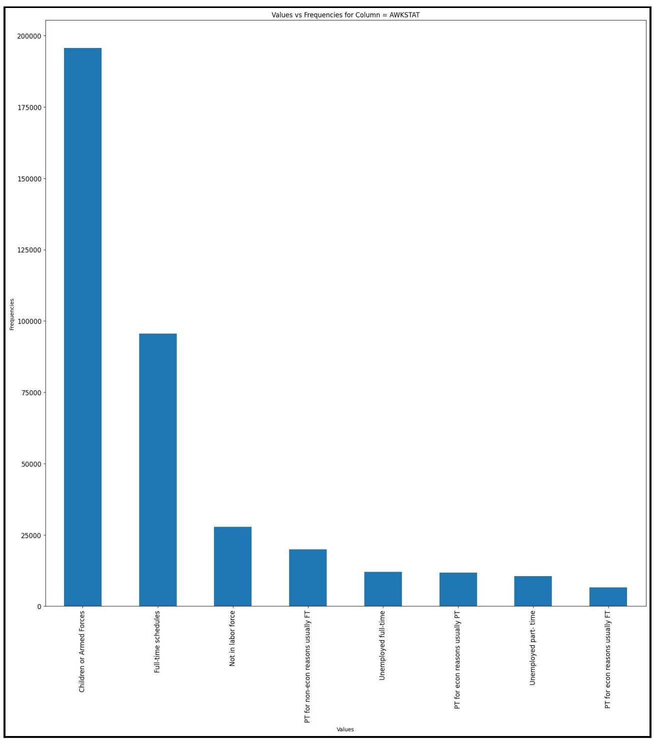

Values vs Frequencies for Column = ASEX
Frequencies
0
500
1000
1500
2000
2500
3000
3500
Female
Male
Values

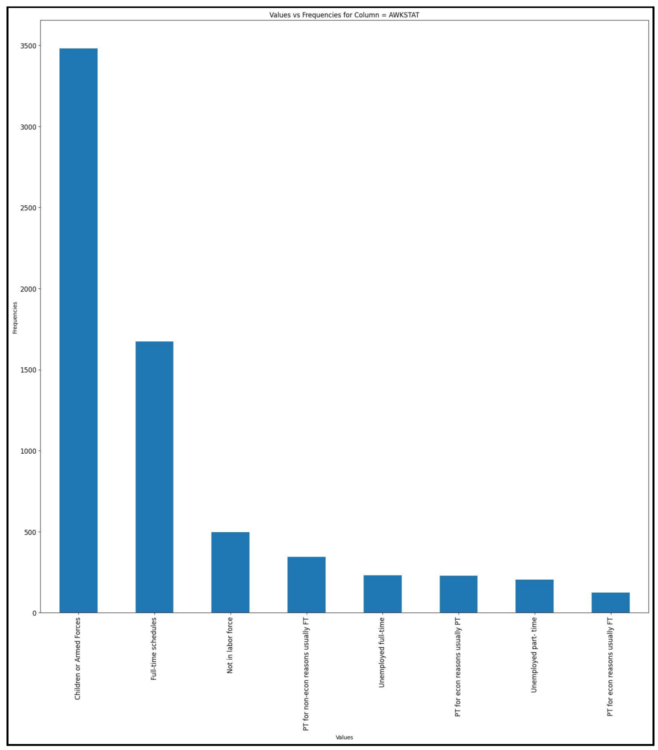

Values vs Frequencies for Column = AUNMEM
Frequencies
0
25000
50000
75000
100000
125000
150000
175000
200000
Not in universe
No
Yes
Values

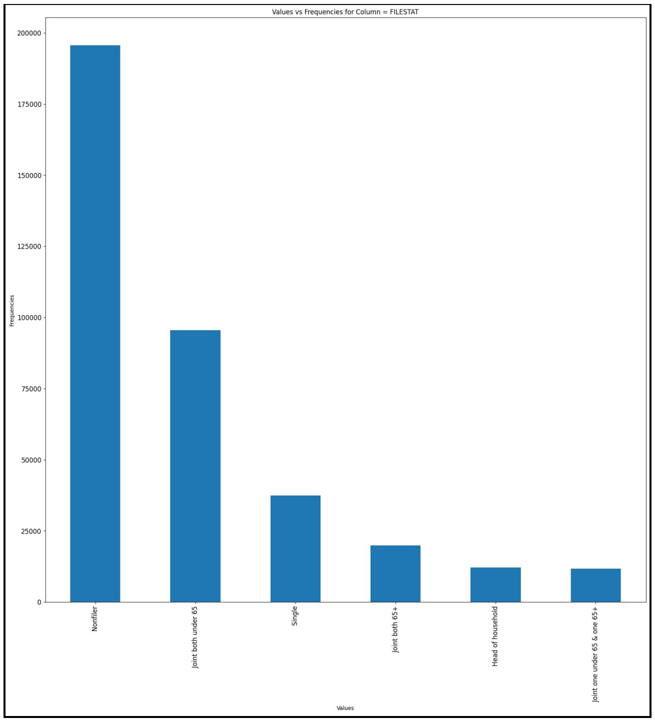

Values vs Frequencies for Column = AUNMEM
Frequencies
0
500
1000
1500
2000
2500
3000
3500
Not in universe
No
Yes
Values

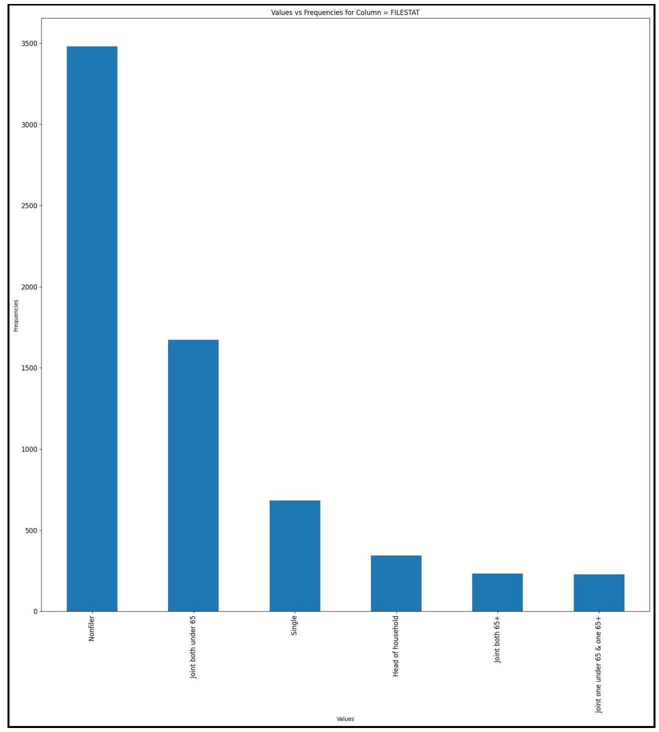

Values vs Frequencies for Column = AUNTYPE
Frequencies
Values
Not in universe
Other job loser
Re-entrant
Job loser - on layoff
Job leaver
New entrant

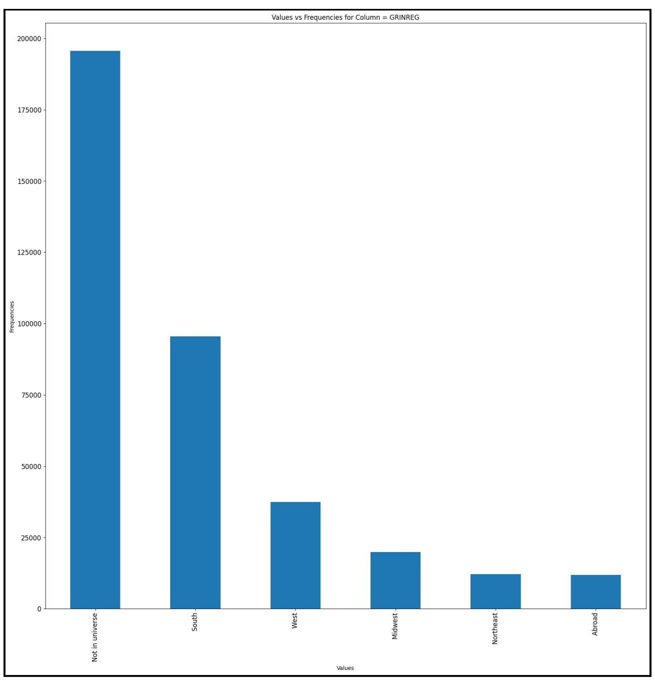

Values vs Frequencies for Column = AUNTYPE
Frequencies
Values
Not in universe
Other job loser
Re-entrant
Job loser - on layoff
New entrant
Job leaver

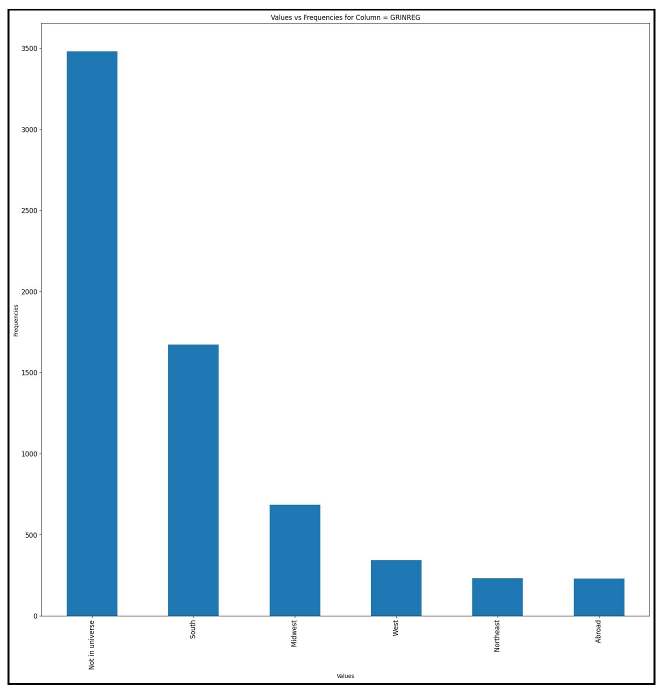

Values vs Frequencies for Column = AWKSTAT
Frequencies
Values
Children or Armed Forces
Full-time schedules
Not in labor force
PT for non-econ reasons usually FT
Unemployed full-time
PT for econ reasons usually PT
Unemployed part- time
PT for econ reasons usually FT

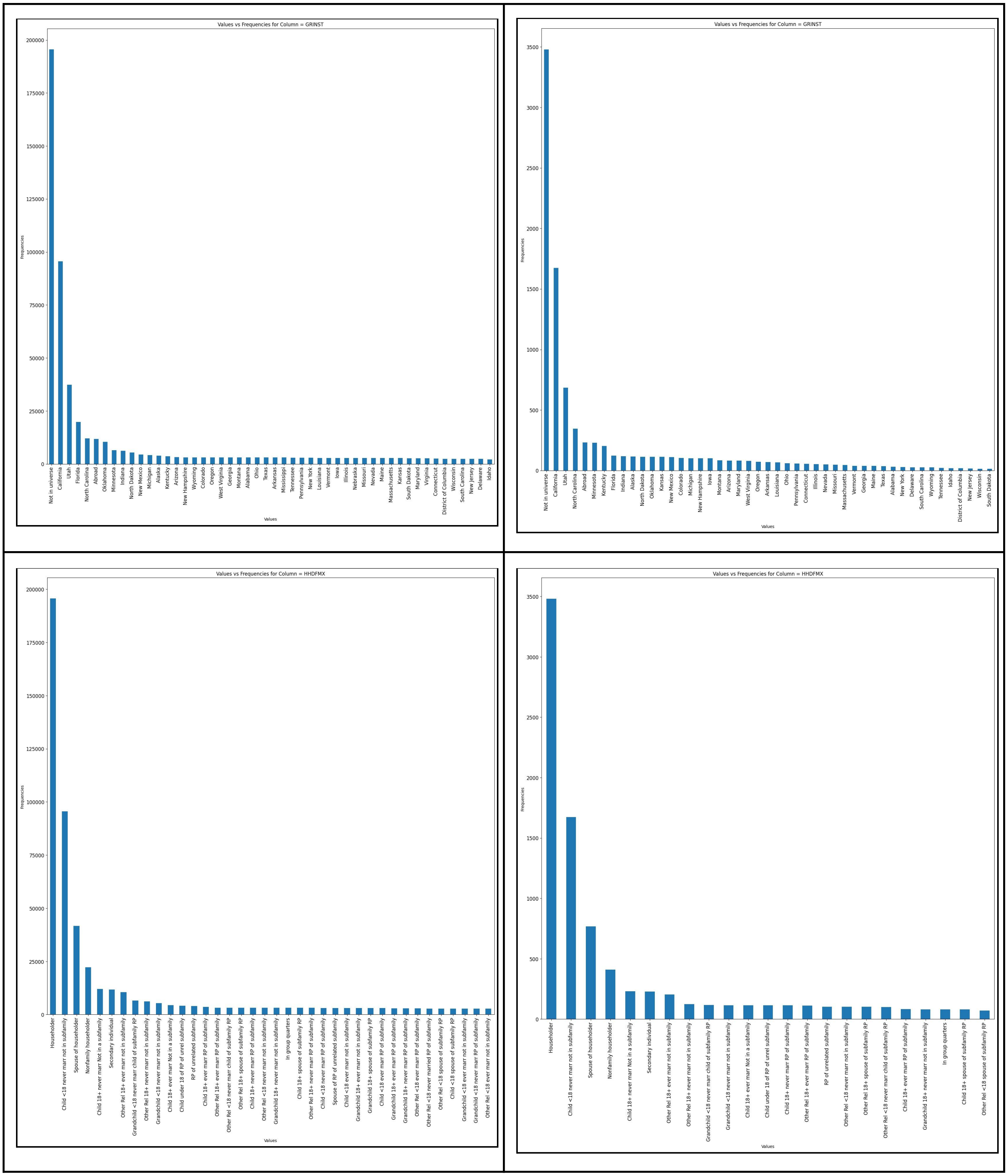

Values vs Frequencies for Column = AWKSTAT
Frequencies
Values
Children or Armed Forces
Full-time schedules
Not in labor force
PT for non-econ reasons usually FT
Unemployed full-time
PT for econ reasons usually PT
Unemployed part- time
PT for econ reasons usually FT


Values vs Frequencies for Column = FILESTAT
Frequencies
200000
175000
150000
125000
100000
75000
50000
25000
0
Nonfiler
Joint both under 65
Single
Joint both 65+
Head of household
Joint one under 65 & one 65+
Values

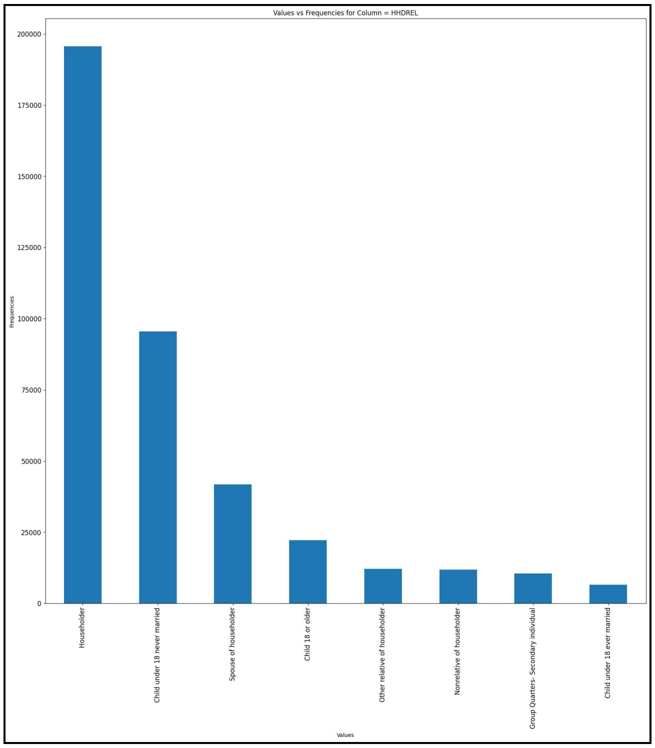

Values vs Frequencies for Column = FILESTAT
Frequencies
3500
3000
2500
2000
1500
1000
500
0
Nonfiler
Joint both under 65
Single
Head of household
Joint both 65+
Joint one under 65 & one 65+
Values

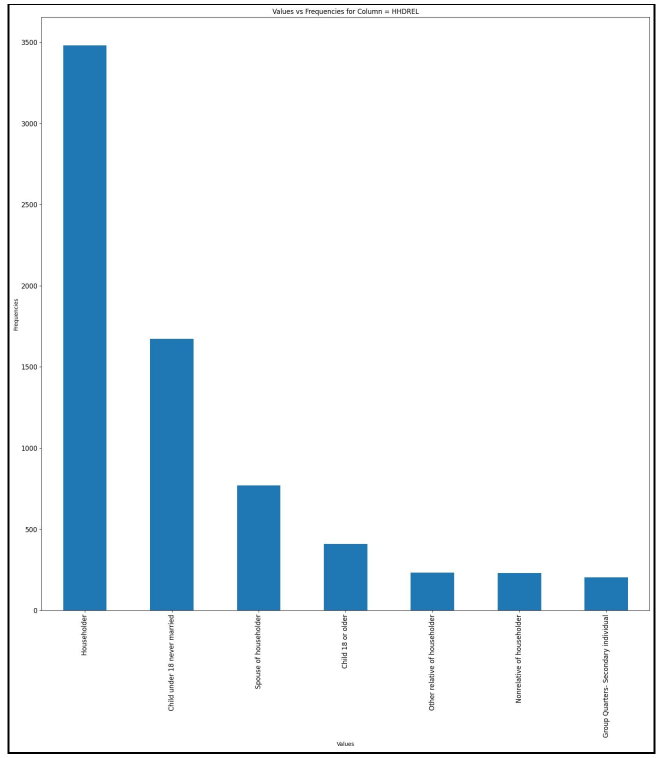

Values vs Frequencies for Column = GRINREG
Frequencies
200000
175000
150000
125000
100000
75000
50000
25000
0
Not in universe
South
West
Midwest
Northeast
Abroad
Values

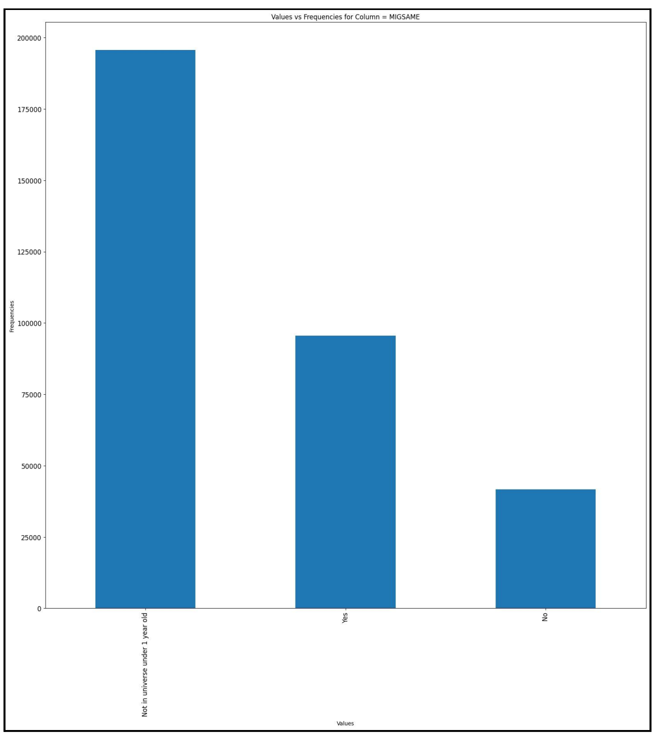

Values vs Frequencies for Column = GRINREG
Frequencies
3500
3000
2500
2000
1500
1000
500
0
Not in universe
South
Midwest
West
Northeast
Abroad
Values

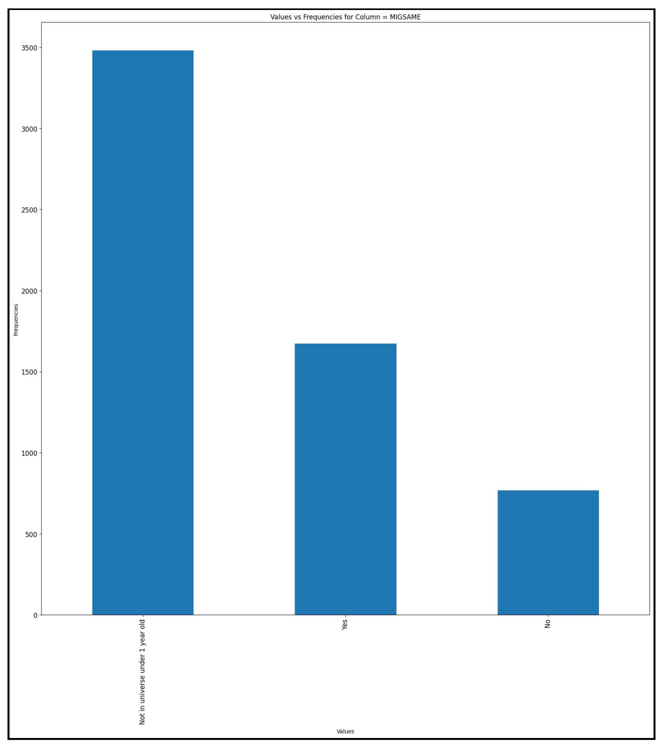


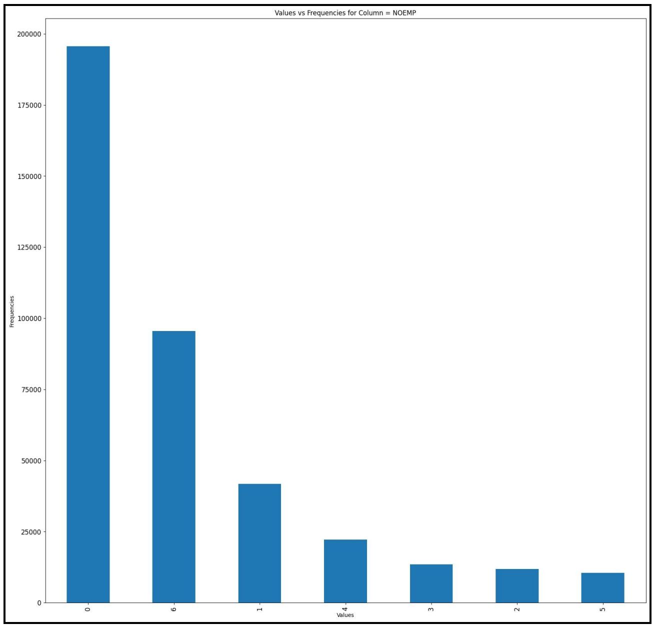
Values vs Frequencies for Column = GRINST
Frequencies
Values
Not in universe
California
Utah
Florida
North Carolina
Abroad
Oklahoma
Minnesota
Indiana
North Dakota
New Mexico
Michigan
Alaska
Kentucky
Arizona
New Hampshire
Wyoming
Colorado
Oregon
West Virginia
Georgia
Montana
Alabama
Ohio
Texas
Arkansas
Mississippi
Tennessee
Pennsylvania
New York
Louisiana
Vermont
Iowa
Illinois
Nebraska
Missouri
Nevada
Maine
Massachusetts
Kansas
South Dakota
Maryland
Virginia
Connecticut
District of Columbia
Wisconsin
South Carolina
New Jersey
Delaware
Idaho
Values vs Frequencies for Column = GRINST
Frequencies
Values
Values vs Frequencies for Column = HHDFMX
Frequencies
Values
Householder
Child <18 never marr not in subfamily
Spouse of householder
Nonfamily householder
Child 18+ never marr Not in a subfamily
Secondary individual
Other Rel 18+ ever marr not in subfamily
Grandchild <18 never marr child of subfamily RP
Other Rel 18+ never marr not in subfamily
Grandchild <18 never marr not in subfamily
Child 18+ ever marr Not in a subfamily
Child under 18 of RP of unrel subfamily
RP of unrelated subfamily
In group quarters
Values vs Frequencies for Column = HHDFMX
Frequencies
Values

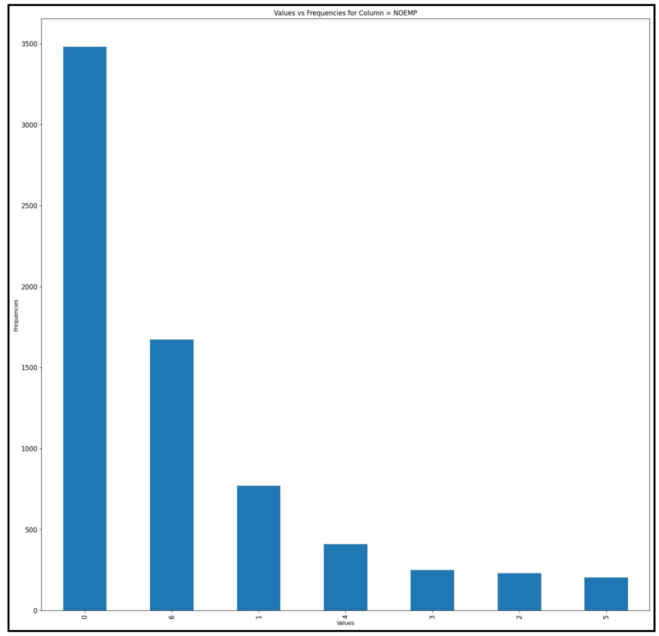

Values vs Frequencies for Column = HHDREL
Frequencies
0
25000
50000
75000
100000
125000
150000
175000
200000
Householder
Child under 18 never married
Spouse of householder
Child 18 or older
Other relative of householder
Nonrelative of householder
Group Quarters- Secondary individual
Child under 18 ever married
Values

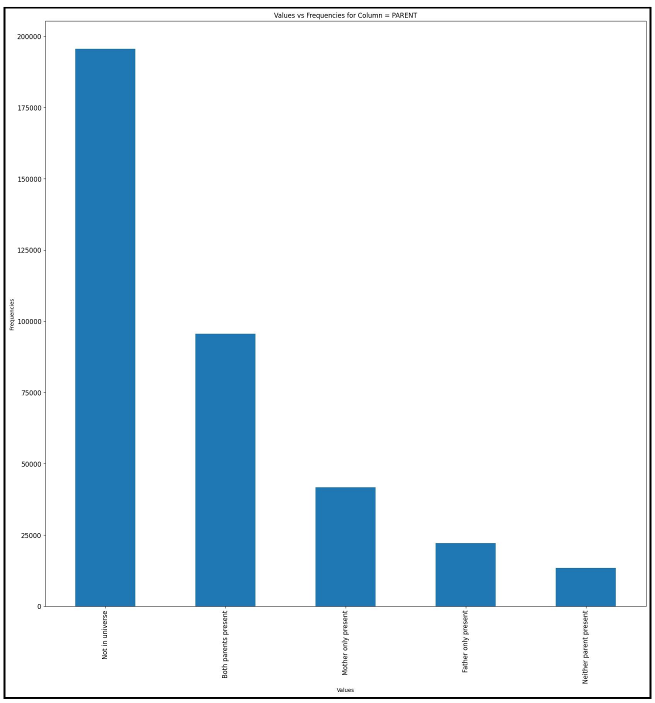

Values vs Frequencies for Column = HHDREL
Frequencies
0
500
1000
1500
2000
2500
3000
3500
Householder
Child under 18 never married
Spouse of householder
Child 18 or older
Other relative of householder
Nonrelative of householder
Group Quarters- Secondary individual
Values

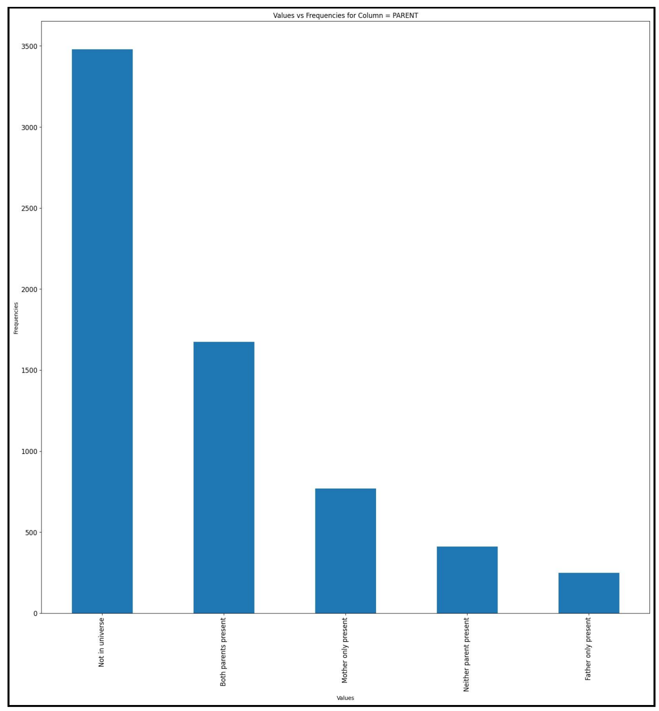

Values vs Frequencies for Column = MIGSAME
Frequencies
0
25000
50000
75000
100000
125000
150000
175000
200000
Not in universe under 1 year old
Yes
No
Values

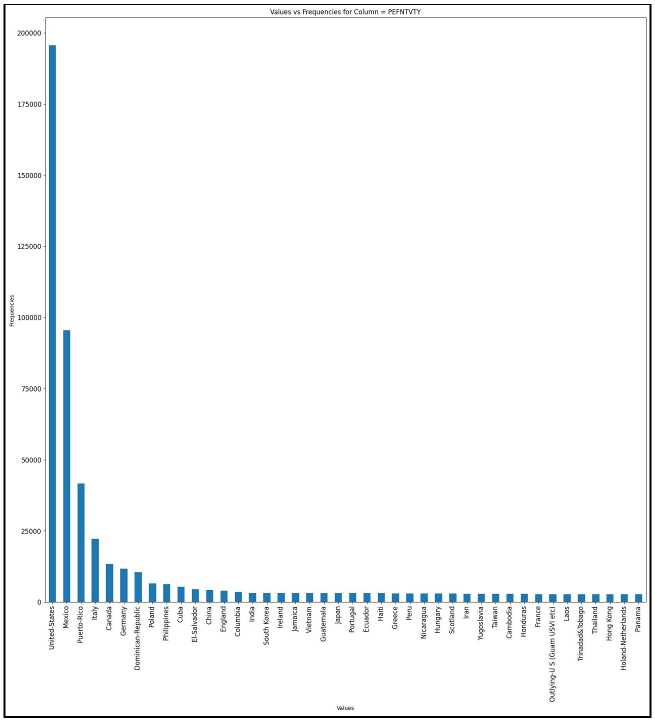

Values vs Frequencies for Column = MIGSAME
Frequencies
0
500
1000
1500
2000
2500
3000
3500
Not in universe under 1 year old
Yes
No
Values

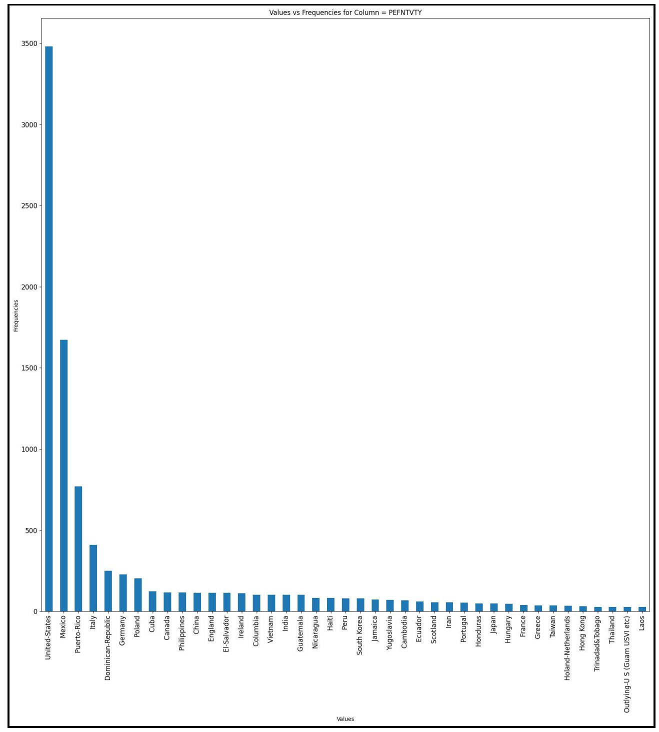

Values vs Frequencies for Column = NOEMP
Frequencies
Values
0
6
1
4
3
2
5
200000
175000
150000
125000
100000
75000
50000
25000
0

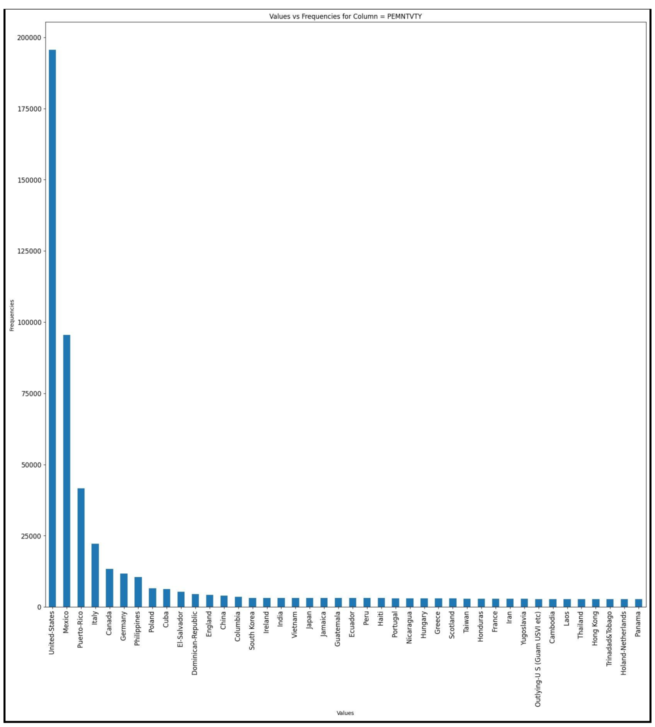

Values vs Frequencies for Column = NOEMP
Frequencies
Values
0
6
1
4
3
2
5
3500
3000
2500
2000
1500
1000
500
0

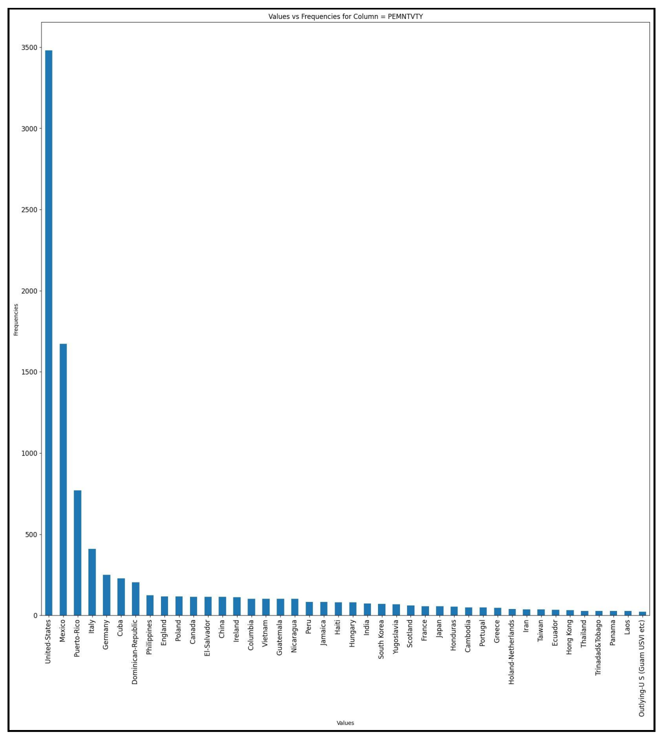

Values vs Frequencies for Column = PARENT
Frequencies
Values
Not in universe
Both parents present
Mother only present
Father only present
Neither parent present
200000
175000
150000
125000
100000
75000
50000
25000
0

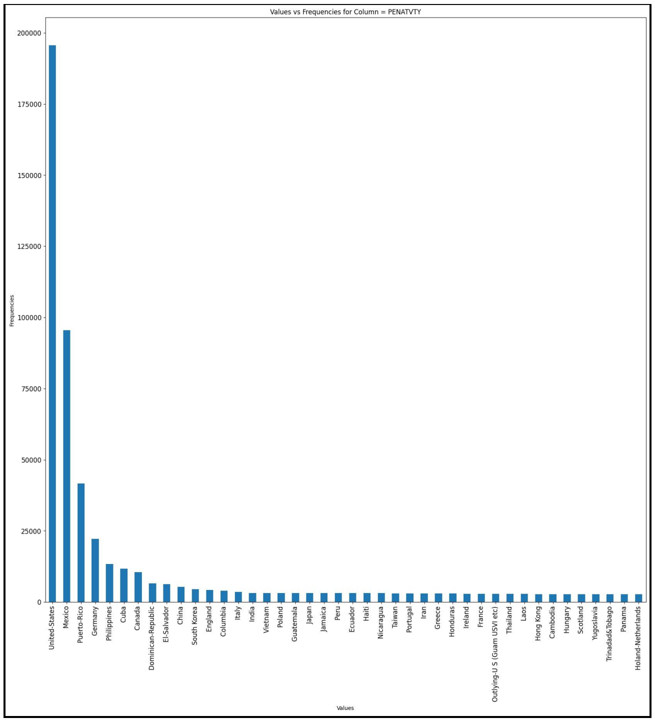

Values vs Frequencies for Column = PARENT
Frequencies
Values
Not in universe
Both parents present
Mother only present
Neither parent present
Father only present
3500
3000
2500
2000
1500
1000
500
0

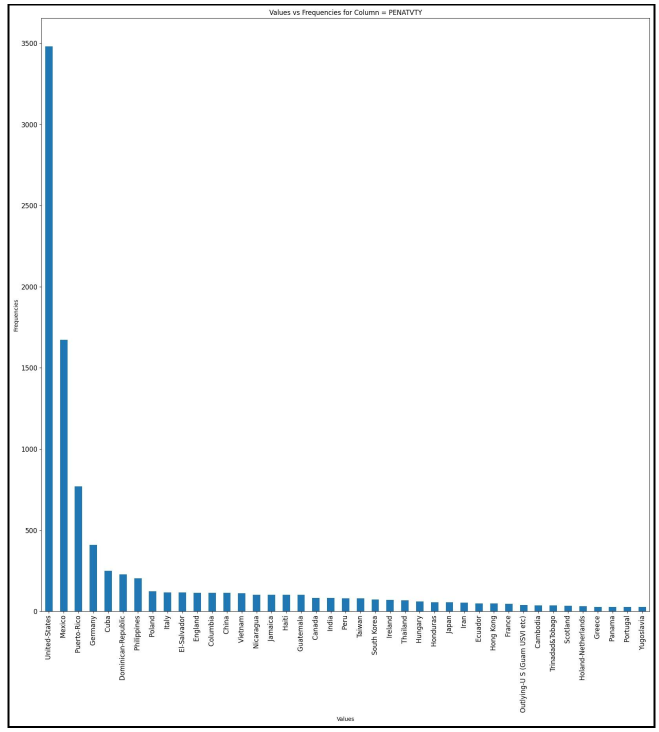

Values vs Frequencies for Column = PEFNTVTY
Frequencies
200000
175000
150000
125000
100000
75000
50000
25000
0
United-States
Mexico
Puerto-Rico
Italy
Canada
Germany
Dominican-Republic
Poland
Philippines
Cuba
El-Salvador
China
England
Columbia
India
South Korea
Ireland
Jamaica
Vietnam
Guatemala
Japan
Portugal
Ecuador
Haiti
Greece
Peru
Nicaragua
Hungary
Scotland
Iran
Yugoslavia
Taiwan
Cambodia
Honduras
France
Outlying-U S (Guam USVI etc)
Laos
Trinadad&Tobago
Thailand
Hong Kong
Holand-Netherlands
Panama
Values

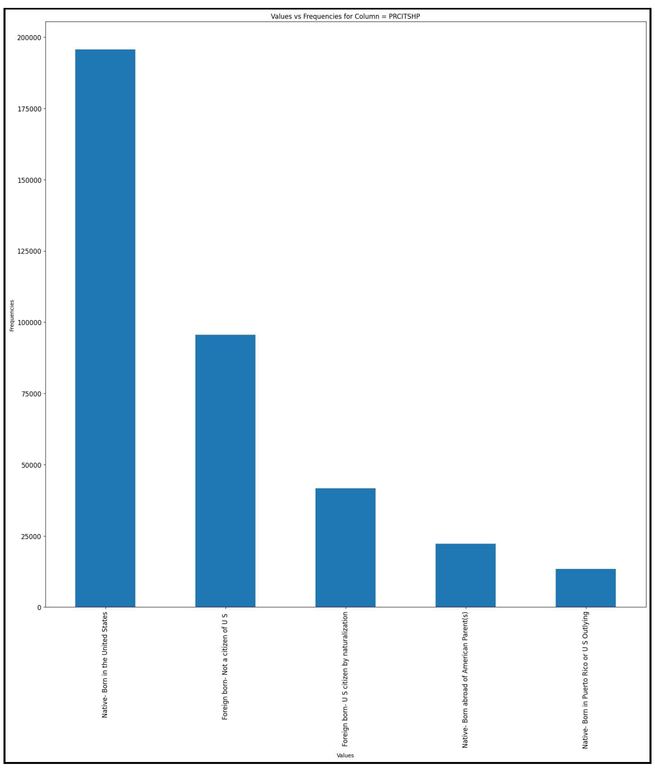

Values vs Frequencies for Column = PEFNTVTY
Frequencies
3500
3000
2500
2000
1500
1000
500
0
United-States
Mexico
Puerto-Rico
Italy
Dominican-Republic
Germany
Poland
Cuba
Canada
Philippines
China
England
El-Salvador
Ireland
Columbia
Vietnam
India
Guatemala
Nicaragua
Haiti
Peru
South Korea
Jamaica
Yugoslavia
Cambodia
Ecuador
Scotland
Iran
Portugal
Honduras
Japan
Hungary
France
Greece
Taiwan
Holand-Netherlands
Hong Kong
Trinadad&Tobago
Thailand
Outlying-U S (Guam USVI etc)
Laos
Values

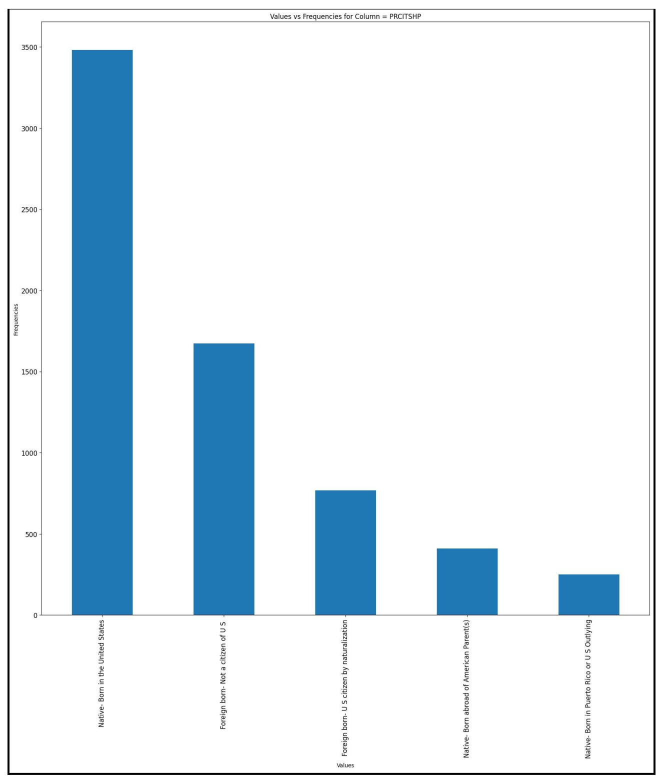

Values vs Frequencies for Column = PEMNTVTY
Frequencies
200000
175000
150000
125000
100000
75000
50000
25000
0
United-States
Mexico
Puerto-Rico
Italy
Canada
Germany
Philippines
Poland
Cuba
El-Salvador
Dominican-Republic
England
China
Columbia
South Korea
Ireland
India
Vietnam
Japan
Jamaica
Guatemala
Ecuador
Peru
Haiti
Portugal
Nicaragua
Hungary
Greece
Scotland
Taiwan
Honduras
France
Iran
Yugoslavia
Outlying-U S (Guam USVI etc)
Cambodia
Laos
Thailand
Hong Kong
Trinadad&Tobago
Holand-Netherlands
Panama
Values

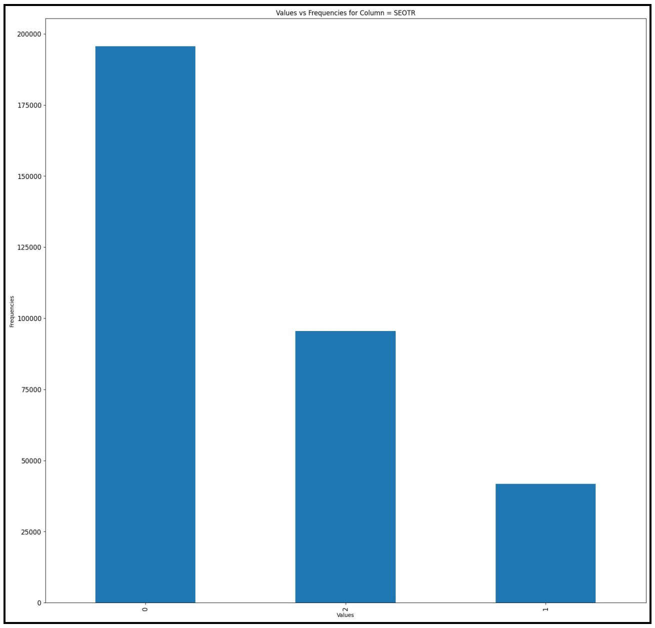

Values vs Frequencies for Column = PEMNTVTY
Frequencies
3500
3000
2500
2000
1500
1000
500
0
United-States
Mexico
Puerto-Rico
Italy
Germany
Cuba
Dominican-Republic
Philippines
England
Poland
Canada
El-Salvador
China
Ireland
Columbia
Vietnam
Guatemala
Nicaragua
Peru
Jamaica
Haiti
Hungary
India
South Korea
Yugoslavia
Scotland
France
Japan
Honduras
Cambodia
Portugal
Greece
Holand-Netherlands
Iran
Taiwan
Ecuador
Hong Kong
Thailand
Trinadad&Tobago
Panama
Laos
Outlying-U S (Guam USVI etc)
Values

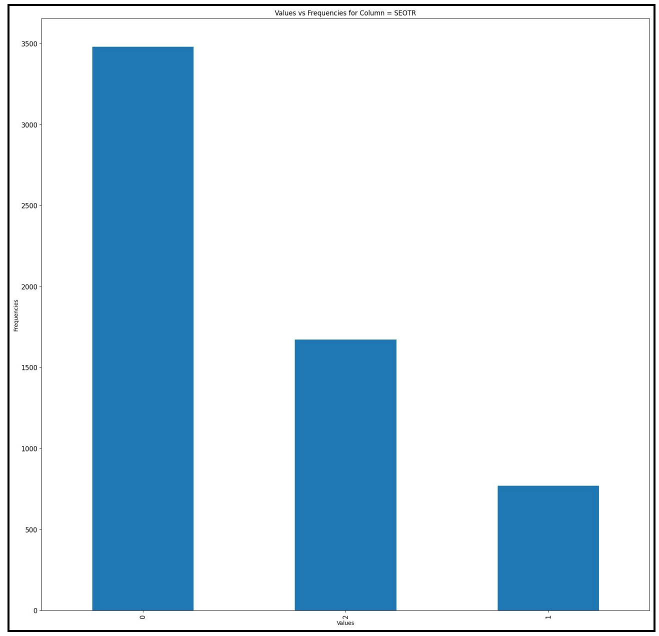

Values vs Frequencies for Column = PENATVTY
Frequencies
Values
0
25000
50000
75000
100000
125000
150000
175000
200000
United-States
Mexico
Puerto-Rico
Germany
Philippines
Cuba
Canada
Dominican-Republic
El-Salvador
China
South Korea
England
Columbia
Italy
India
Vietnam
Poland
Guatemala
Japan
Jamaica
Peru
Ecuador
Haiti
Nicaragua
Taiwan
Portugal
Iran
Greece
Honduras
Ireland
France
Outlying-U S (Guam USVI etc)
Thailand
Laos
Hong Kong
Cambodia
Hungary
Scotland
Yugoslavia
Trinadad&Tobago
Panama
Holand-Netherlands

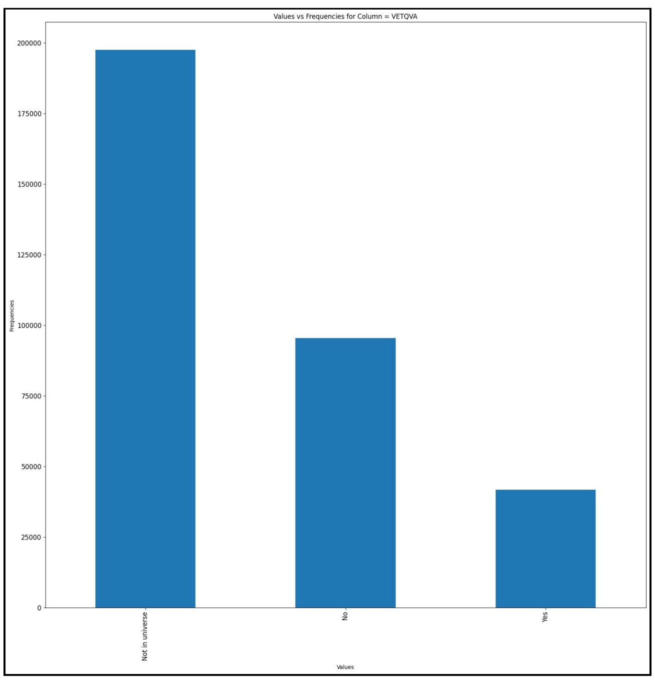

Values vs Frequencies for Column = PENATVTY
Frequencies
Values
0
500
1000
1500
2000
2500
3000
3500
United-States
Mexico
Puerto-Rico
Germany
Cuba
Dominican-Republic
Philippines
Poland
Italy
El-Salvador
England
Columbia
China
Vietnam
Nicaragua
Jamaica
Haiti
Guatemala
Canada
India
Peru
Taiwan
South Korea
Ireland
Thailand
Hungary
Honduras
Japan
Iran
Ecuador
Hong Kong
France
Outlying-U S (Guam USVI etc)
Cambodia
Trinadad&Tobago
Scotland
Holand-Netherlands
Greece
Panama
Portugal
Yugoslavia

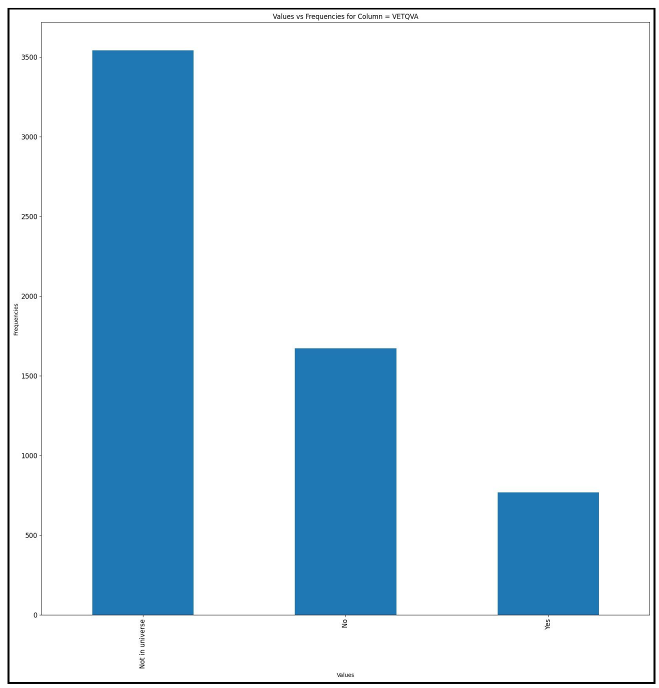

Values vs Frequencies for Column = PRCITSHP
Frequencies
Values
0
25000
50000
75000
100000
125000
150000
175000
200000
Native- Born in the United States
Foreign born- Not a citizen of U S
Foreign born- U S citizen by naturalization
Native- Born abroad of American Parent(s)
Native- Born in Puerto Rico or U S Outlying

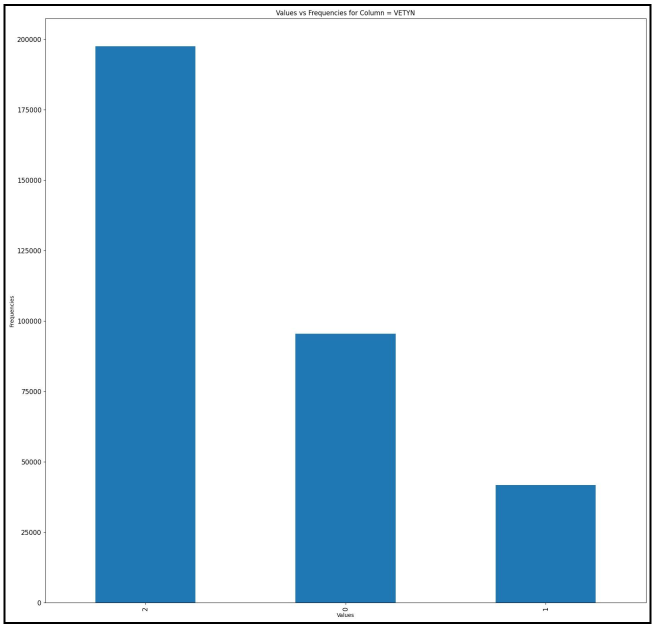

Values vs Frequencies for Column = PRCITSHP
Frequencies
Values
0
500
1000
1500
2000
2500
3000
3500
Native- Born in the United States
Foreign born- Not a citizen of U S
Foreign born- U S citizen by naturalization
Native- Born abroad of American Parent(s)
Native- Born in Puerto Rico or U S Outlying

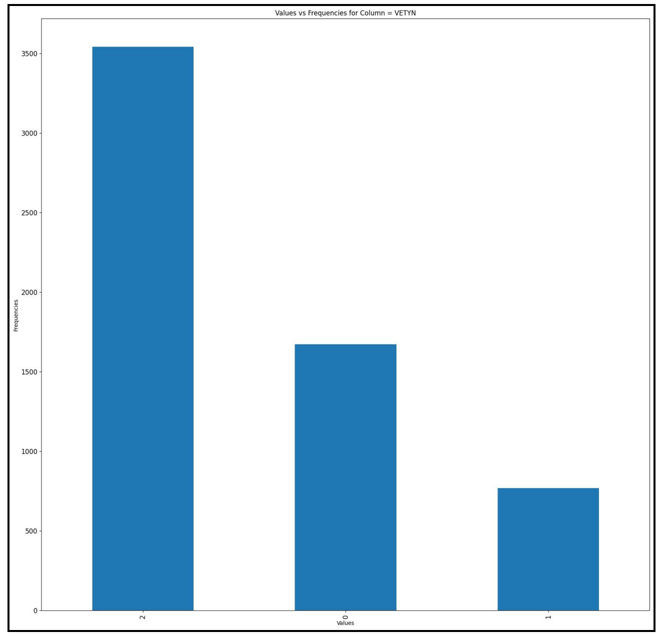

Values vs Frequencies for Column = SEOTR
Frequencies
200000
175000
150000
125000
100000
75000
50000
25000
0
0
2
1
Values

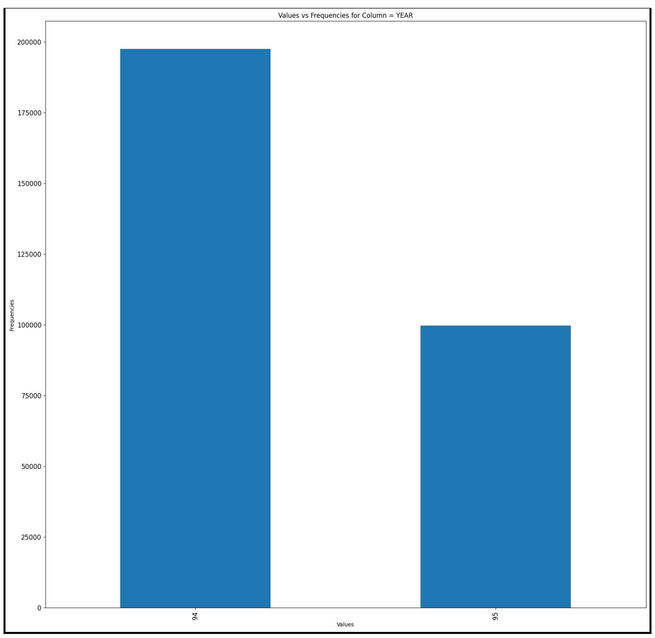

Values vs Frequencies for Column = SEOTR
Frequencies
3500
3000
2500
2000
1500
1000
500
0
0
2
1
Values

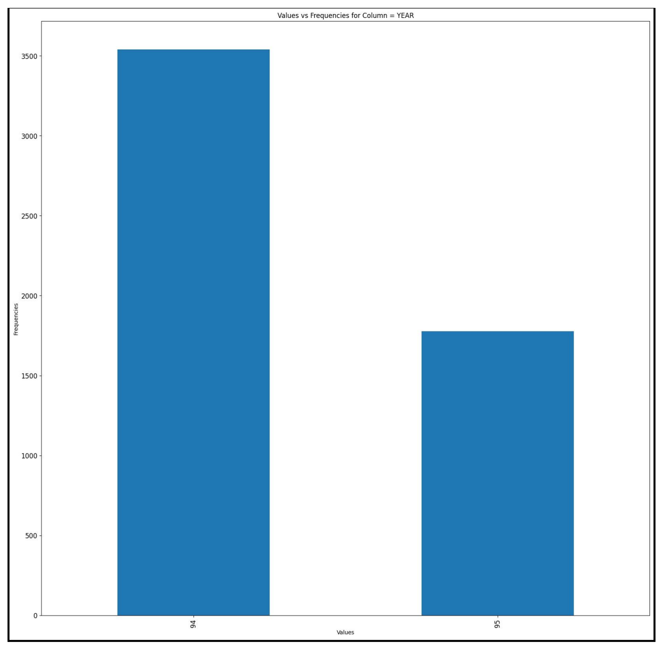

Values vs Frequencies for Column = VETQVA
Frequencies
200000
175000
150000
125000
100000
75000
50000
25000
0
Not in universe
No
Yes
Values


Values vs Frequencies for Column = VETQVA
Frequencies
3500
3000
2500
2000
1500
1000
500
0
Not in universe
No
Yes
Values



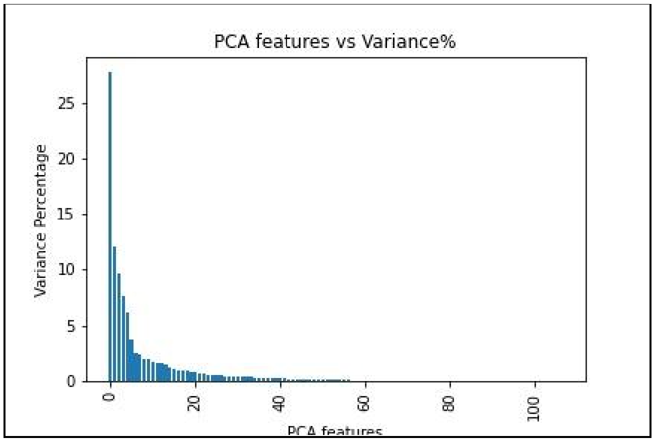
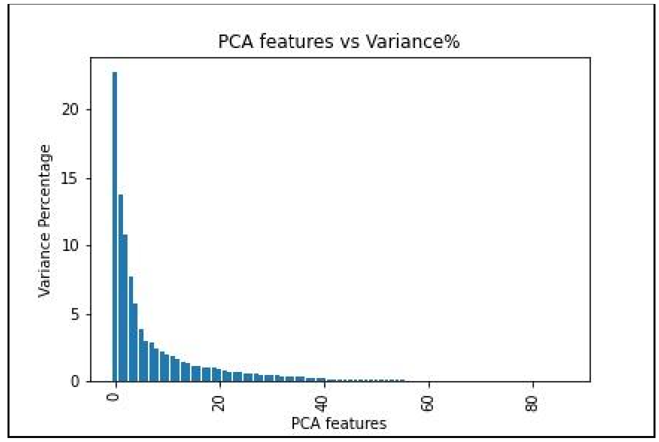




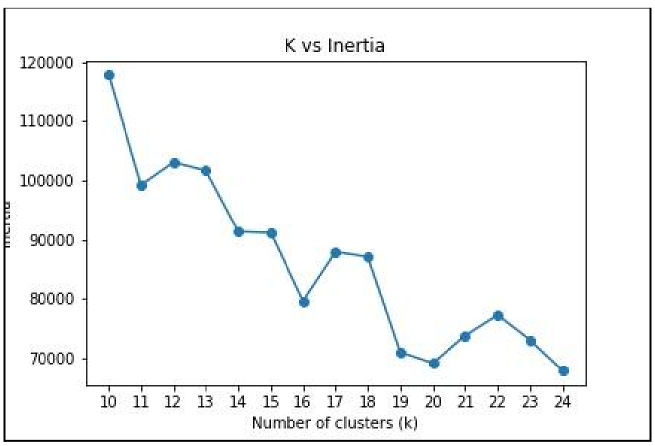
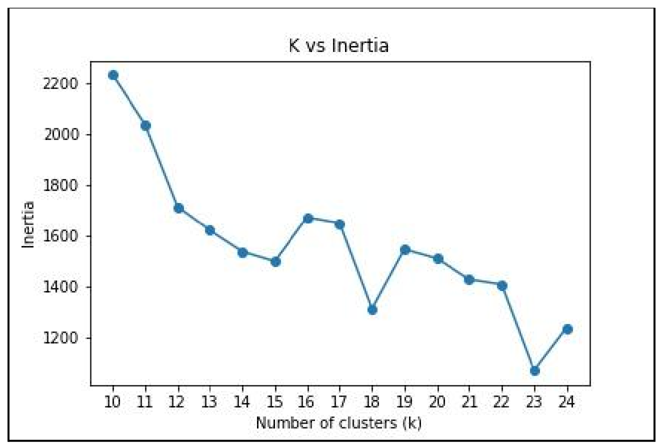


PCA Features vs Variance Percentage

population.csv	**more_than_50K.csv**

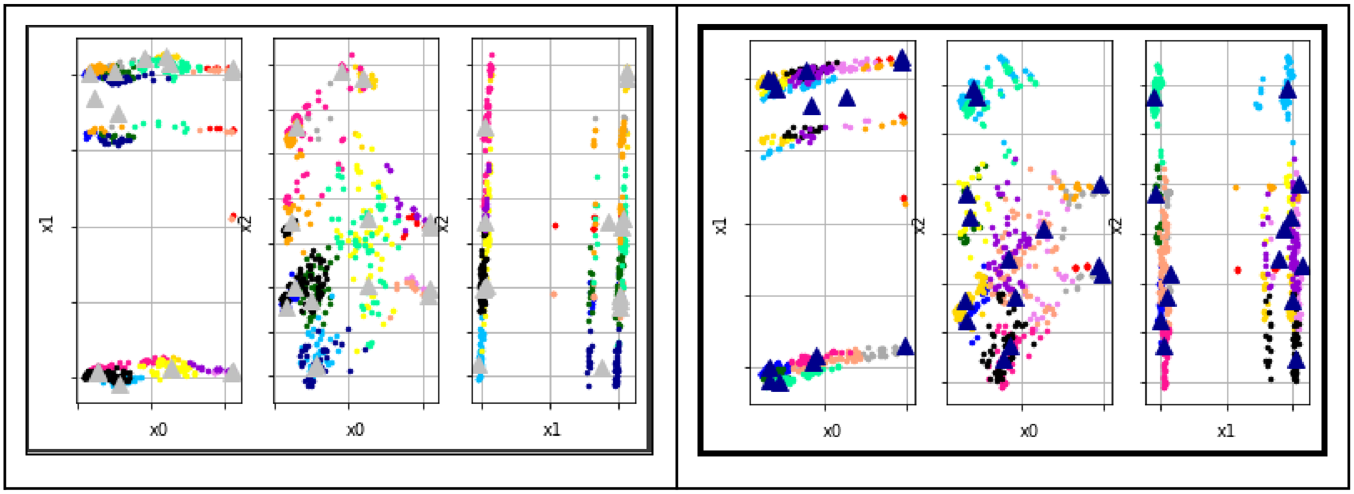
<table>
<tr>
<td>
</td>
<td>
</td>
</tr>
</table>

k vs Inertia for [10, 24]

<table>
<tr>
<th>population.csv</th>
<th>more_than_50K.csv</th>
</tr>
<tr>
<td>
</td>
<td>
</td>
</tr>
</table>

Clusters for k = 20

population.csv	more_than_50K.csv



Question 6.2, 6.3, 6.4 :

From the generated visuals over-represented clusters are
1 3 4 8 9 10 13 14 15 17 for General Population data. While all clusters are overrepresented in More than 50k data.

Question 2 and 3 :

Question - 2

a) Yes, we need a non-negativity constraint on ε.
To always hold the condition $y^{(i)}(w^T x^{(i)} + b) \geq 1 - \varepsilon_i$,
ε_i must be non negative.
$\therefore$ ε_i should be defined as $\max(\varepsilon_i, 0)$ to put the non-negativity constraint on ε_i

b) Lagrangian of above Optimization problem -

$$L(w, b, \varepsilon, \alpha) = \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{m}\varepsilon_i^2 + \sum \alpha_i (1 - y_i [w^T x + b] - \varepsilon_i)$$

c) Required Dual is —
Using Karush Kuhn Tucker condition,

$$\frac{\partial L}{\partial w} = 0$$

$$\Rightarrow w - \sum_{i=1}^{m} \alpha_i y_i x_i = 0$$

$$\Rightarrow \boxed{w = \sum_{i=1}^{m} \alpha_i y_i x_i} \quad --- \quad (1)$$

$$\frac{\partial L}{\partial b} = 0$$

$$\Rightarrow \boxed{\sum \alpha_i y_i = 0} \quad --- \quad (2)$$

$$\frac{\partial L}{\partial \xi} = 0$$

$$\Rightarrow C\xi - \alpha_i = 0$$

$$\Rightarrow \boxed{\xi = \frac{\alpha_i}{C}} \quad \text{---} \; (3)$$

From (1), (2) & (3)

Dual is

$$\max \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \, x^{(i)^T} x^{(j)}$$

satisfying $\sum \alpha_i y_i = 0$ and $0 \leq \alpha_i \leq C$

Question 3

a) $$f(x) = \sum_{i=1}^{m} \alpha_i y^{(i)} K(x^{(i)}, x) + b$$

For $\alpha_i = 1$ & $b = 0$ for $1 \leq i \leq m$, we have

$$|f(x^i) - y^i| = \left| \sum_{j=1}^{m} y^j K(x^j, x) - y^i \right|$$

$$= \left| \sum_{j=1}^{m} y^j e^{-\frac{\|x^j - x^i\|^2}{\tau^2}} - y^i \right| \qquad \left[\text{Given, } K(x, z) = e^{-\frac{\|x - z\|^2}{\tau}}\right]$$

$$= \left| y^i + \sum_{j \neq i} y^j e^{-\frac{\|x^j - x^i\|^2}{\tau^2}} - y^i \right|$$

$$= \left| \sum_{j \neq i} y^j e^{-\frac{\|x^j - x^i\|^2}{\tau^2}} \right|$$

$$\leq \sum_{j \neq i} \left| y^j e^{-\frac{\|x^j - x^i\|^2}{\tau^2}} \right| \qquad \left[\text{As } |\Sigma a| \leq \Sigma |a|\right]$$

$$= \sum_{j \neq i} |y^j| e^{-\frac{\|x^j - x^i\|^2}{\tau^2}}$$

$$= \sum_{j \neq i} e^{-\frac{\varepsilon^2}{\tau^2}} \qquad \left[|\gamma^j| e^{-\|x^j - x^i\|^2} = e^{-\varepsilon^2} \right]$$

$$= (m-1)\, e^{-\frac{\varepsilon^2}{\tau^2}}$$

Now $(m-1)\, e^{-\frac{\varepsilon^2}{\tau^2}} < 1$

$$\Rightarrow \cancel{\frac{2\varepsilon^2}{\tau^2}} \ln\left[(m-1) \cdot e^{-\frac{\varepsilon^2}{\tau^2}}\right] < \ln 1$$

$$\Rightarrow \ln(m-1) - \left(\frac{\varepsilon}{\tau}\right)^2 \ln e < 0$$

$$\Rightarrow \boxed{\tau < \frac{\varepsilon}{\ln(m-1)}}$$

b) To have a zero training error we would need ~~all~~ relative weights to be zero. Relative weights are controlled by α and $\|w\|$. Now, weights can never by zero $\left[x^T w_i = 0 \text{ does not make sense} \right]$ In turn the relative weight controlled by α and $\|w\|$ won't be zero, i.e. we can have close to zero training error but not zero exactly.