Paper1_checkpoint_2

April 15, 2021

```
import pandas as pd import numpy as np
df_2017 = pd.read_csv('hmda_2017_nationwide_all-records_labels.csv', dtype = {'msamd': str, 'state_code': str, 'county_code': str, 'census_tract_number': str, 'denial_reason_1':str, 'denial_reason_2': str, 'denial_reason_3': str})
df_2017.head()
```

as_of_year respondent_id agency_name \ 0 2017 75-2921540 Department of Housing and Urban Development 1 2017 0000504713 Consumer Financial Protection Bureau 2 2017 7810600004 Department of Housing and Urban Development 3 2017 42-1739728 Department of Housing and Urban Development 4 2017 42-1739728 Department of Housing and Urban Development

agency_abbr agency_code loan_type_name loan_type \ 0 HUD 7 Conventional 1 1 CFPB 9 Conventional 1 2 HUD 7 Conventional 1 3 HUD 7 Conventional 1 4 HUD 7 Conventional 1

property_type_name property_type \ 0 One-to-four family dwelling (other than manufa… 1 1 One-to-four family dwelling (other than manufa… 1 2 One-to-four family dwelling (other than manufa… 1 3 One-to-four family dwelling (other than manufa… 1 4 One-to-four family dwelling (other than manufa… 1

loan_purpose_name loan_purpose owner_occupancy_name \ 0 Refinancing 3 Owner-occupied as a principal dwelling 1 Refinancing 3 Owner-occupied as a principal dwelling 2 Refinancing 3 Owner-occupied as a principal dwelling 3 Refinancing 3 Not owner-occupied as a principal dwelling 4 Refinancing 3 Owner-occupied as a principal dwelling

owner_occupancy loan_amount_000s preapproval_name preapproval \ 0 1 53.0 Not applicable 3 1 1 168.0 Not applicable 3 2 1 103.0 Not applicable 3 3 2 88.0 Not applicable 3 4 1 90.0 Not applicable 3

action_taken_name action_taken \ 0 Application withdrawn by applicant 4 1 Application denied by financial institution 3 2 File closed for incompleteness 5 3 Loan originated 1 4 Application withdrawn by applicant 4

msamd_name msamd state_name \ 0 Philadelphia - PA 37964 Pennsylvania 1 Spokane, Spokane Valley - WA 44060 Washington 2 Salt Lake City - UT 41620 Utah 3 Springfield - MO 44180 Missouri 4 Chicago, Naperville, Arlington Heights - IL 16974 Illinois

state_abbr state_code county_name county_code census_tract_number \ 0 PA 42 Philadelphia County 101 0173.00 1 WA 53 Spokane County 63 0127.01 2 UT 49 Salt Lake County 35 1136.00 3 MO 29 Greene County 77 0011.00 4 IL 17 Cook County 31 0306.04

applicant_ethnicity_name applicant_ethnicity \ 0 Not Hispanic or Latino 2 1 Not Hispanic or Latino 2 2 Information not provided by applicant in mail,… 3 3 Not Hispanic or Latino 2 4 Not Hispanic or Latino 2

co_applicant_ethnicity_name co_applicant_ethnicity \ 0 No co-applicant 5 1 Not Hispanic or Latino 2 2 Information not provided by applicant in mail,… 3 3 No co-applicant 5 4 No co-applicant 5

applicant_race_name_1 applicant_race_1 \ 0 Black or African American 3 1 White 5 2 Information not provided by applicant in mail,... 6 3 White 5 4 White 5

applicant_race_name_2 applicant_race_2 applicant_race_name_3 \ 0 NaN NaN NaN 1 NaN NaN NaN 2 NaN NaN NaN 3 NaN NaN NaN 4 NaN NaN NaN

applicant_race_3 applicant_race_name_4 applicant_race_4 \ 0 NaN NaN NaN 1 NaN NaN NaN 2 NaN NaN NaN 3 NaN NaN NaN 4 NaN NaN NaN

applicant_race_name_5 applicant_race_5 \ 0 NaN NaN 1 NaN NaN 2 NaN NaN 3 NaN NaN 4 NaN NaN

co_applicant_race_name_1 co_applicant_race_1 \ 0 No co-applicant 8 1 White 5 2 Information not provided by applicant in mail,... 6 3 No co-applicant 8 4 No co-applicant 8

co_applicant_race_name_2 co_applicant_race_2 co_applicant_race_name_3 \ 0 NaN NaN NaN 1 NaN NaN NaN 2 NaN NaN NaN 3 NaN NaN NaN 4 NaN NaN NaN

co_applicant_race_3 co_applicant_race_name_4 co_applicant_race_4 \ 0 NaN NaN NaN 1 NaN NaN NaN 2 NaN NaN NaN 3 NaN NaN NaN 4 NaN NaN NaN

co_applicant_race_name_5 co_applicant_race_5 applicant_sex_name \ 0 NaN NaN Male 1 NaN NaN Male 2 NaN NaN Male 3 NaN NaN Female 4 NaN NaN Male

applicant_sex co_applicant_sex_name co_applicant_sex \ 0 1 No co-applicant 5 1 1 Female 2 2 1 Female 2 3 2 No co-applicant 5 4 1 No co-applicant 5

applicant_income_000s purchaser_type_name \ 0 12.0 Loan was not originated or was not sold in cal... 1 60.0 Loan was not originated or was not sold in cal... 2 50.0 Loan was not originated or was not sold in cal... 3 53.0 Freddie Mac (FHLMC) 4 29.0 Loan was not originated or was not sold in cal...

purchaser_type denial_reason_name_1 denial_reason_1 denial_reason_name_2 \ 0 0 NaN NaN NaN 1 0 Credit history 3.0 NaN 2 0 NaN NaN NaN 3 3 NaN NaN NaN 4 0 NaN NaN NaN

denial_reason_2 denial_reason_name_3 denial_reason_3 rate_spread \ 0 NaN NaN NaN NaN 1 NaN NaN NaN NaN 2 NaN NaN NaN NaN 3 NaN NaN NaN NaN 4 NaN NaN NaN NaN

hoepa_status_name hoepa_status lien_status_name lien_status \ 0 Not a HOEPA loan 2 Secured by a first lien 1 1 Not a HOEPA loan 2 Secured by a first lien 1 2 Not a HOEPA loan 2 Secured by a first lien 1 3 Not a HOEPA loan 2 Secured by a first lien 1 4 Not a HOEPA loan 2 Secured by a first lien 1

edit_status_name edit_status sequence_number population \ 0 NaN NaN NaN 3202.0 1 NaN NaN NaN 3733.0 2 NaN NaN NaN 5498.0 3 NaN NaN NaN 3566.0 4 NaN NaN NaN 2910.0

minority_population hud_median_family_income tract_to_msamd_income \ 0 97.279999 57400.0 47.540001 1 4.580000 63900.0 86.239998 2 37.919998 75400.0 63.939999 3 11.830000 55200.0 74.290001 4 48.660000 77500.0 79.250000

number_of_owner_occupied_units number_of_1_to_4_family_units \ 0 710.0 1314.0 1 861.0 1241.0 2 1270.0 1658.0 3 573.0 1261.0 4 599.0 26.0

application_date_indicator 0 NaN 1 NaN 2 NaN 3 NaN 4 NaN

df_2017.columns

Index(['as_of_year', 'respondent_id', 'agency_name', 'agency_abbr', 'agency_code', 'loan_type_name', 'loan_type', 'property_type_name', 'property_type', 'loan_purpose_name', 'loan_purpose', 'owner_occupancy_name', 'owner_occupancy', 'loan_amount_000s', 'preapproval_name', 'preapproval', 'action_taken_name', 'action_taken', 'msamd_name', 'msamd', 'state_name', 'state_abbr', 'state_code', 'county_name', 'county_code', 'census_tract_number', 'applicant_ethnicity_name', 'applicant_ethnicity', 'co_applicant_ethnicity_name', 'co_applicant_ethnicity', 'applicant_race_name_1', 'applicant_race_1', 'applicant_race_name_2',

'applicant_race_2', 'applicant_race_name_3', 'applicant_race_3', 'applicant_race_name_4', 'applicant_race_4', 'applicant_race_name_5', 'applicant_race_5', 'co_applicant_race_name_1', 'co_applicant_race_1', 'co_applicant_race_name_2', 'co_applicant_race_2', 'co_applicant_race_name_3', 'co_applicant_race_3', 'co_applicant_race_name_4', 'co_applicant_race_4', 'co_applicant_race_name_5', 'co_applicant_race_5', 'applicant_sex_name', 'applicant_sex', 'co_applicant_sex_name', 'co_applicant_sex', 'applicant_income_000s', 'purchaser_type_name', 'purchaser_type', 'denial_reason_name_1', 'denial_reason_1', 'denial_reason_name_2', 'denial_reason_2', 'denial_reason_name_3', 'denial_reason_3', 'rate_spread', 'hoepa_status_name', 'hoepa_status', 'lien_status_name', 'lien_status', 'edit_status_name', 'edit_status', 'sequence_number', 'population', 'minority_population', 'hud_median_family_income', 'tract_to_msamd_income', 'number_of_owner_occupied_units', 'number_of_1_to_4_family_units', 'application_date_indicator'], dtype='object')

```
pd.set_option('display.max_columns', None)
df_2017.head()
```

as_of_year respondent_id agency_name \ 0 2017 75-2921540 Department of Housing and Urban Development 1 2017 0000504713 Consumer Financial Protection Bureau 2 2017 7810600004 Department of Housing and Urban Development 3 2017 42-1739728 Department of Housing and Urban Development 4 2017 42-1739728 Department of Housing and Urban Development

agency_abbr agency_code loan_type_name loan_type \ 0 HUD 7 Conventional 1 1 CFPB 9 Conventional 1 2 HUD 7 Conventional 1 3 HUD 7 Conventional 1 4 HUD 7 Conventional 1

property_type_name property_type \ 0 One-to-four family dwelling (other than manufa… 1 1 One-to-four family dwelling (other than manufa… 1 2 One-to-four family dwelling (other than manufa… 1 3 One-to-four family dwelling (other than manufa… 1 4 One-to-four family dwelling (other than manufa… 1

loan_purpose_name loan_purpose owner_occupancy_name \ 0 Refinancing 3 Owner-occupied as a principal dwelling 1 Refinancing 3 Owner-occupied as a principal dwelling 2 Refinancing 3 Owner-occupied as a principal dwelling 3 Refinancing 3 Not owner-occupied as a principal dwelling 4 Refinancing 3 Owner-occupied as a principal dwelling

owner_occupancy loan_amount_000s preapproval_name preapproval \ 0 1 53.0 Not applicable 3 1 1 168.0 Not applicable 3 2 1 103.0 Not applicable 3 3 2 88.0 Not applicable 3 4 1 90.0 Not applicable 3

action_taken_name action_taken \ 0 Application withdrawn by applicant 4 1 Application denied by financial institution 3 2 File closed for incompleteness 5 3 Loan originated 1 4 Application withdrawn by applicant 4

msamd_name msamd state_name \ 0 Philadelphia - PA 37964 Pennsylvania 1 Spokane, Spokane Valley - WA 44060 Washington 2 Salt Lake City - UT 41620 Utah 3 Springfield - MO 44180 Missouri 4 Chicago, Naperville, Arlington Heights - IL 16974 Illinois

state_abbr state_code county_name county_code census_tract_number \ 0 PA 42 Philadelphia County 101 0173.00 1 WA 53 Spokane County 63 0127.01 2 UT 49 Salt Lake County 35 1136.00 3 MO 29 Greene County 77 0011.00 4 IL 17 Cook County 31 0306.04

applicant_ethnicity_name applicant_ethnicity \ 0 Not Hispanic or Latino 2 1 Not Hispanic or Latino 2 2 Information not provided by applicant in mail,… 3 3 Not Hispanic or Latino 2 4 Not Hispanic or Latino 2

co_applicant_ethnicity_name co_applicant_ethnicity \ 0 No co-applicant 5 1 Not Hispanic or Latino 2 2 Information not provided by applicant in mail,… 3 3 No co-applicant 5 4 No co-applicant 5

applicant_race_name_1 applicant_race_1 \ 0 Black or African American 3 1 White 5 2 Information not provided by applicant in mail,… 6 3 White 5 4 White 5

applicant_race_name_2 applicant_race_2 applicant_race_name_3 \ 0 NaN NaN NaN 1 NaN NaN NaN 2 NaN NaN NaN 3 NaN NaN NaN 4 NaN NaN NaN

applicant_race_3 applicant_race_name_4 applicant_race_4 \ 0 NaN NaN NaN 1 NaN NaN NaN 2 NaN NaN NaN 3 NaN NaN NaN 4 NaN NaN NaN

applicant_race_name_5 applicant_race_5 \ 0 NaN NaN 1 NaN NaN 2 NaN NaN 3 NaN NaN 4 NaN NaN

co_applicant_race_name_1 co_applicant_race_1 \ 0 No co-applicant 8 1 White 5 2 Information not provided by applicant in mail,... 6 3 No co-applicant 8 4 No co-applicant 8

co_applicant_race_name_2 co_applicant_race_2 co_applicant_race_name_3 \ 0 NaN NaN NaN 1 NaN NaN NaN 2 NaN NaN NaN 3 NaN NaN NaN 4 NaN NaN NaN

co_applicant_race_3 co_applicant_race_name_4 co_applicant_race_4 \ 0 NaN NaN NaN 1 NaN NaN NaN 2 NaN NaN NaN 3 NaN NaN NaN 4 NaN NaN NaN

co_applicant_race_name_5 co_applicant_race_5 applicant_sex_name \ 0 NaN NaN Male 1 NaN NaN Male 2 NaN NaN Male 3 NaN NaN Female 4 NaN NaN Male

applicant_sex co_applicant_sex_name co_applicant_sex \ 0 1 No co-applicant 5 1 1 Female 2 2 1 Female 2 3 2 No co-applicant 5 4 1 No co-applicant 5

applicant_income_000s purchaser_type_name \ 0 12.0 Loan was not originated or was not sold in cal... 1 60.0 Loan was not originated or was not sold in cal... 2 50.0 Loan was not originated or was not sold in cal... 3 53.0 Freddie Mac (FHLMC) 4 29.0 Loan was not originated or was not sold in cal...

purchaser_type denial_reason_name_1 denial_reason_1 denial_reason_name_2 \ 0 0 NaN NaN NaN 1 0 Credit history 3.0 NaN 2 0 NaN NaN NaN 3 3 NaN NaN NaN 4 0 NaN NaN NaN

denial_reason_2 denial_reason_name_3 denial_reason_3 rate_spread \ 0 NaN NaN NaN NaN 1 NaN NaN NaN NaN 2 NaN NaN NaN NaN 3 NaN NaN NaN NaN 4 NaN NaN NaN NaN

hoepa_status_name hoepa_status lien_status_name lien_status \ 0 Not a HOEPA loan 2 Secured by a first lien 1 1 Not a HOEPA loan 2 Secured by a first lien 1 2 Not a HOEPA loan 2 Secured by a first lien 1 3 Not a HOEPA loan 2 Secured by a first lien 1 4 Not a HOEPA loan 2 Secured by a first lien 1

edit_status_name edit_status sequence_number population \ 0 NaN NaN NaN 3202.0 1 NaN NaN NaN 3733.0 2 NaN NaN NaN 5498.0 3 NaN NaN NaN 3566.0 4 NaN NaN NaN 2910.0

minority_population hud_median_family_income tract_to_msamd_income \ 0 97.279999 57400.0 47.540001 1 4.580000 63900.0 86.239998 2 37.919998 75400.0 63.939999 3 11.830000 55200.0 74.290001 4 48.660000 77500.0 79.250000

number_of_owner_occupied_units number_of_1_to_4_family_units \ 0 710.0 1314.0 1 861.0 1241.0 2 1270.0 1658.0 3 573.0 1261.0 4 599.0 26.0

application_date_indicator 0 NaN 1 NaN 2 NaN 3 NaN 4 NaN

0.1 Regarding Application Status Based on Race

```
df_race_action = df_2017[['action_taken_name','action_taken','applicant_race_name_1','applicant_race_1']]
df_race_action
```

action_taken_name action_taken \ 0 Application withdrawn by applicant 4 1 Application denied by financial institution 3 2 File closed for incompleteness 5 3 Loan originated 1 4 Application withdrawn by applicant 4 ... ... 14285491 Loan originated 1 14285492 Application denied by financial institution 3 14285493 Loan originated 1 14285494 Application denied by financial institution 3 14285495 Application denied by financial institution 3

applicant_race_name_1 applicant_race_1 0 Black or African American 3 1 White 5 2 Information not provided by applicant in mail,... 6 3 White 5 4 White 5 ... ... 14285491 Information not provided by applicant in mail,... 6 14285492 Information not provided by applicant in mail,... 6 14285493 Information not provided by applicant in mail,... 6 14285494 White 5 14285495 White 5

[14285496 rows x 4 columns]

white_action = df_race_action[df_race_action['applicant_race_1'] == 5] white_action

action_taken_name action_taken \ 1 Application denied by financial institution 3 3 Loan originated 1 4 Application withdrawn by applicant 4 5 Application withdrawn by applicant 4 9 Loan originated 1 ... ... 14285483 File closed for incompleteness 5 14285487 Application approved but not accepted 2 14285489 Application denied by financial institution 3 14285494 Application denied by financial institution 3 14285495 Application denied by financial institution 3

applicant_race_name_1 applicant_race_1 1 White 5 3 White 5 4 White 5 5 White 5 9 White 5 ... ... 14285483 White 5 14285487 White 5 14285489 White 5 14285494 White 5 14285495 White 5

[9267426 rows x 4 columns]

white_action['action_taken'].value_counts(normalize=**True**) * 100

1 58.993101 3 13.936836 4 12.444372 6 6.344793 5 4.211396 2 3.129952 7 0.666690 8 0.272859 Name: action_taken, dtype: float64

black_action = df_race_action[df_race_action['applicant_race_1'] == 3] black_action

action_taken_name action_taken \ 0 Application withdrawn by applicant 4 10 Application denied by financial institution 3 25 Application denied by financial institution 3 30 Application denied by financial institution 3 32 Application withdrawn by applicant 4 ... ... 14285457 Application withdrawn by applicant 4 14285461 Application denied by financial institution 3 14285475 Application withdrawn by applicant 4 14285476 Loan originated 1 14285485 Application withdrawn by applicant 4

applicant_race_name_1 applicant_race_1 0 Black or African American 3 10 Black or African American 3 25 Black or African American 3 30 Black or African American 3 32 Black or African American 3 ... ... 14285457 Black or African American 3 14285461 Black or African American 3 14285475 Black or African American 3 14285476 Black or African American 3 14285485 Black or African American 3

[1002556 rows x 4 columns]

black_action['action_taken'].value_counts(normalize=**True**) * 100

1 45.334724 3 24.862152 4 14.197212 5 5.989391 6 3.897338 2 3.360810 7 2.195089 8 0.163283 Name: action_taken, dtype: float64

na_action = df_race_action[df_race_action['applicant_race_1'] == 1] na_action

action_taken_name action_taken \ 33 File closed for incompleteness 5 97 Loan originated 1 232 Loan originated 1 420 Loan originated 1 566 Loan originated 1 ... ... 14285211 Application denied by financial institution 3 14285219 Application denied by financial institution 3 14285246 Application approved but not accepted 2 14285338 Application withdrawn by applicant 4 14285436 File closed for incompleteness 5

applicant_race_name_1 applicant_race_1 33 American Indian or Alaska Native 1 97 American Indian or Alaska Native 1 232 American Indian or Alaska Native 1 420 American Indian or Alaska Native 1 566 American Indian or Alaska Native 1 ... ... 14285211 American Indian or Alaska Native 1 14285219 American Indian or Alaska Native 1 14285246 American Indian or Alaska Native 1 14285338 American Indian or Alaska Native 1 14285436 American Indian or Alaska Native 1

[109588 rows x 4 columns]

na_action['action_taken'].value_counts(normalize=**True**) * 100

1 44.667299 3 25.251852 4 14.837391 5 6.781764 6 3.629959 2 3.164580 7 1.488302 8 0.178852 Name: action_taken, dtype: float64

asian_action = df_race_action[df_race_action['applicant_race_1'] == 2] asian_action['action_taken'].value_counts(normalize=True) * 100

1 58.277401 4 13.691896 3 12.536359 6 6.712867 5 4.321741 2 3.233973 7 0.732019 8 0.493744 Name: action_taken, dtype: float64

pa_action = df_race_action[df_race_action['applicant_race_1'] == 4] pa_action['action_taken'].value_counts(normalize=True) * 100

1 51.913531 3 18.328900 4 14.479814 5 5.507649 6 4.997713 2 3.015569 7 1.597574 8 0.159249 Name: action_taken, dtype: float64

0.2 Regarding Application Status Based on Sex of Applicant and Co-Applicant

df_action_gender = df_2017[['applicant_sex_name','applicant_sex','co_applicant_sex_name','co_applicant_sex','ac

m_f = df_action_gender[(df_action_gender['applicant_sex']==1) & (df_action_gender['co_applicant_sex'] == 2)] m_f['action_taken'].value_counts(normalize=True) * 100

1 61.082976 4 12.395166 3 11.823354 6 6.582495 5 4.188526 2 3.055464 7 0.595727 8 0.276292 Name: action_taken, dtype: float64

f_m = df_action_gender[(df_action_gender['applicant_sex']==2) & (df_action_gender['co_applicant_sex'] == 1)] f_m['action_taken'].value_counts(normalize=True) * 100

1 58.772253 3 14.716190 4 12.203630 6 5.869638 5 4.144491 2 3.055725 7 0.913898 8 0.324175 Name: action_taken, dtype: float64

m_m = df_action_gender[(df_action_gender['applicant_sex']==1) & (df_action_gender['co_applicant_sex'] == 1)] m_m['action_taken'].value_counts(normalize=True) * 100

1 58.003196 3 15.100793 4 11.925589 6 7.169793 5 3.974963 2 2.903945 7 0.735274 8 0.186447 Name: action_taken, dtype: float64

f_f = df_action_gender[(df_action_gender['applicant_sex']==2) & (df_action_gender['co_applicant_sex'] == 2)] f_f['action_taken'].value_counts(normalize=True) * 100

1 55.149407 3 17.625607 4 12.118420 6 6.512960 5 4.525981 2 2.909834 7 0.953566 8 0.204226 Name: action_taken, dtype: float64

0.3 Looking Into Reasons for Denial By Race

df_denials = df_2017[['action_taken_name','action_taken','applicant_race_name_1','applicant_race_1', 'denial_reason_name_1', 'denial_reason_1', 'denial_reason_name_2', 'denial_reason_2', 'denial_reason_name_3', 'denial_reason_3',]]

df_denials.head()

action_taken_name action_taken \ 0 Application withdrawn by applicant 4 1 Application denied by financial institution 3 2 File closed for incompleteness 5 3 Loan originated 1 4 Application withdrawn by applicant 4

applicant_race_name_1 applicant_race_1 \ 0 Black or African American 3 1 White 5 2 Information not provided by applicant in mail,... 6 3 White 5 4 White 5

denial_reason_name_1 denial_reason_1 denial_reason_name_2 denial_reason_2 \ 0 NaN NaN NaN NaN 1 Credit history 3 NaN NaN 2 NaN NaN NaN NaN 3 NaN NaN NaN NaN 4 NaN NaN NaN NaN

denial_reason_name_3 denial_reason_3 0 NaN NaN 1 NaN NaN 2 NaN NaN 3 NaN NaN 4 NaN NaN

na_denial = df_denials[df_denials['applicant_race_1'] == 1] na_counts = na_denial[['denial_reason_1','denial_reason_2','denial_reason_3']] count = pd.Series(na_counts.squeeze().values.ravel()).value_counts() pd.DataFrame({'Denial Reason': count.index, 'Count':count.values, 'Percentage':(count/count.sum()).values})

Denial Reason Count Percentage 0 3 8328 0.369542 1 1 5798 0.257277 2 9 2381 0.105653 3 4 2147 0.095270 4 7 1740 0.077210 5 5 936 0.041534 6 6 777 0.034478 7 2 384 0.017039 8 8 45 0.001997

as_denial = df_denials[df_denials['applicant_race_1'] == 2] as_counts = as_denial[['denial_reason_1','denial_reason_2','denial_reason_3']] count = pd.Series(as_counts.squeeze().values.ravel()).value_counts() pd.DataFrame({'Denial Reason': count.index, 'Count':count.values, 'Percentage':(count/count.sum()).values})

Denial Reason Count Percentage 0 1 26872 0.318019 1 3 15163 0.179448 2 4 9839 0.116441 3 9 9304 0.110109 4 7 9289 0.109932 5 6 6302 0.074582 6 5 5325 0.063019 7 2 2253 0.026663 8 8 151 0.001787

bl_denial = df_denials[df_denials['applicant_race_1'] == 3] bl_counts = bl_denial[['denial_reason_1','denial_reason_2','denial_reason_3']] count = pd.Series(bl_counts.squeeze().values.ravel()).value_counts() pd.DataFrame({'Denial Reason': count.index, 'Count':count.values, 'Percentage':(count/count.sum()).values})

Denial Reason Count Percentage 0 3 86281 0.373391 1 1 58779 0.254373 2 9 25282 0.109411 3 4 21887 0.094719 4 7 15457 0.066892 5 5 12601 0.054532 6 6 6923 0.029960 7 2 3472 0.015025 8 8 392 0.001696

pa_denial = df_denials[df_denials['applicant_race_1'] == 4] pa_counts = pa_denial[['denial_reason_1','denial_reason_2','denial_reason_3']] count = pd.Series(pa_counts.squeeze().values.ravel()).value_counts() pd.DataFrame({'Denial Reason': count.index, 'Count':count.values, 'Percentage':(count/count.sum()).values})

Denial Reason Count Percentage 0 3 3395 0.313221 1 1 3065 0.282775 2 9 1233 0.113756 3 4 1057 0.097518 4 7 887 0.081834 5 5 576 0.053141 6 6 411 0.037919 7 2 202 0.018636 8 8 13 0.001199

wh_denial = df_denials[df_denials['applicant_race_1'] == 5] wh_counts = wh_denial[['denial_reason_1','denial_reason_2','denial_reason_3']] count = pd.Series(wh_counts.squeeze().values.ravel()).value_counts() pd.DataFrame({'Denial Reason': count.index, 'Count':count.values, 'Percentage':(count/count.sum()).values})

Denial Reason Count Percentage 0 3 307209 0.277187 1 1 285249 0.257373 2 4 161464 0.145685 3 9 115495 0.104208 4 7 112956 0.101917 5 5 51627 0.046582 6 6 48620 0.043869 7 2 23504 0.021207 8 8 2185 0.001971