Please add rows as necessary

Data set name: **compas-scores-updated.csv**

Generally: This data set is the cleaned version of compas-scores.csv. Includes columns for violent vs non-violent crimes and time spent in jail.

|  |  |
| --- | --- |
| **Attribute** | **Explanation** |
| sex | Gender (M or F) |
| age | Whole number representing years |
| age\_cat | Age category: either ‘Less than 25’, ‘25 - 45’, ‘Greater than 45’ |
| race | African-American, Caucasian, Hispanic, Native American, Asian, Other |
| juv\_fel\_count | Juvenille felony count: how many felonies received before 18 |
| juv\_misd\_count | Juvenille misdemeanor count: how many misdemeanors received before 18 |
| juv\_other\_count | Juvenille other count: how many ‘other’s received before 18. Either formal warnings or they do not know |
| priors\_count | How many total prior offenses. Does not include juvenile ones above |
| days\_b\_screening\_arrest | Days between charge date and arrest |
| c\_days\_from\_compas | Days charge was from compas |
| c\_jail\_in | Timestamp of when person went to jail. YYYY-MM-DD hh:mm:ss. Blank if the person did not go to jail |
| c\_jail\_out | Timestamp of when person got out of jail. YYYY-MM-DD hh:mm:ss. Blank if the person did not go to jail |
| c\_charge\_degree | Degree of charge: F for felony, M for misdemeanor, O for ‘other’. Other is infraction or unknown. |
| c\_charge\_desc | Description of the charge.ex. ‘Assault’, ‘Battery’, etc. |
| is\_recid | 1 if person recidivised, 0 if they did not, -1 if we don’t know |
| decile\_score | 1-10 score of how likely person is predicted by Compas to recidivise (higher number means higher chance to recidivise) |
| score\_text | How likely compas algorithm thinks person is to recidivise. ‘Low’, ‘Medium’ and ‘High’ |
| c\_charge\_violent | Whether charge is violent. V for violent, N for non-violent, O for other. Other means either there charge description was blank or ‘arrest case no charge’ |
| c\_time\_in\_jail | How long a person spent in jail in days (decimals too) |

Data Set name: **allegations\_202007271729**

Generally: This data set is complaints about police officers, referred to here as MOS. It has demographics on both the officer and the person who filed the complaint (complainant).

|  |  |
| --- | --- |
| **Attribute** | **Explanation** |
| unique\_mos\_id | MOS meaning “member of service”, unique identifier |
| first\_name | MOS first name |
| last\_name | MOS last name |
| command\_now | *Current commander/superior ??* |
| shield\_no | MOS shield number |
| complaint\_id | Unique identifier for the complaint. It is possible that there are multiple complaint\_id’s for one incident, for example one might be filed as “discourtesy” and another as “offensive language” but they are logged as separate entries. |
| month\_received | Month the complaint was received |
| year\_received | Year the complaint was received |
| month\_closed | Month the complaint was closed |
| year\_closed | Year the complaint was closed |
| command\_at\_incident | *Commander/superior at time of incident ??* |
| rank\_abbrev\_incident | MOS rank at time of incident (abbreviated). |
| rank\_abbrev\_now | MOS rank now (abbreviated). |
| rank\_now | MOS rank now |
| rank\_incident | MOS rank at time of incident |
| mos\_ethnicity | MOS ethnicity |
| mos\_gender | MOS gender |
| mos\_age\_incident | MOS age when incident occurred |
| complainant\_ethnicity | Complainant being the person who filed a complaint, complainant ethnicity |
| complainant\_gender | Gender of complainant |
| complainant\_age\_incident | Age of complainant at time of incident |
| fado\_type | FADO is an Acronym categorizing type of abuse. It stands for “excessive Force, Abuse of authority, Discourtesy, and Offensive language” |
| allegation | Complainant allegation against police |
| precinct | Precint number where complaint filed ?? |
| contact\_reason | Reason the police were contacted complaint filed? |
| outcome\_description | Description of outcome. (Could use more research, I’m a bit confused on who this applies to). |
| board\_disposition | Board ruling on whether or not officer was in the wrong and consequences. (Could use more research on who the board is, but I think it is within the police force) |

Data set name: **compas-scores-raw.csv**

|  |  |
| --- | --- |
| **Attribute** | **Explanation** |
| Person\_ID | Unique identification number for individuals taking the test |
| Assessment\_ID | Unique identification number for the test the individual took (i.e. their test number) |
| Case\_ID | Unique identification number for the individual’s case |
| Agency\_Text | Where was the test administered |
| LastName | Individual’s last name |
| FirstName | Individual’s first name |
| MiddleName | Individual’s middle name |
| Sex\_Code\_Text | Individual’s sex as a string |
| Ethnic\_Code\_Text | Individual’s ethnicity as a string |
| DateOfBirth | Individual’s date of birth |
| ScaleSet\_ID | ID for the type of test the individual is taking |
| ScaleSet | Type of test the individual is taking as a string |
| AssessmentReason | Reason for the assessment being given |
| Language | Language the assessment is given in |
| LegalStatus | Where in the legal process the test is being given |
| CustodyStatus | Individual’s custody status |
| MaritalStatus | Individual’s marital status |
| Screening\_Date | Date the test was administered |
| RecSupervisionLevel | Number for the recommended supervision level while in custody |
| RecSupervisionLevelText | Text that correlates and deciphers the recommended supervision level |
| Scale\_ID | How much the given test weighs in the overall calculation for the individual |
| DisplayText | Type of test given (i.e. what is the test looking for) |
| RawScore | Raw score for the specific test |
| DecileScore | Adjusted score for the specific test |
| ScoreText | Text related to the decile score that ranks the level of the test |
| AssessmentType | What type of assessment was given (i.e. new person, repeat offender, etc.) |
| IsCompleted | If assessment was completed |
| IsDeleted | If assessment was deleted |