EXPLORATION

<p>For this project, we utilized data from the National Longitudinal Survey of Youth 1997, sponsored by the U.S. Bureau of Labor Statistics. This study was dedicated to tracking the labor market and other life experiences of American men and women.</p>

The 1997 study was chosen over the 1979 study because it is closer (and therefore more relevant) to the ages of the majority of our classmates. The 1997 study also included variables that were of interest, like subjects’ cohabitation.

We utilized Pandas, MatPlotLib, and Plotly to explore the data and determine the needs and focus of our model. This allowed us to visualize the potential relationships between certain variables, as you will see <a href=’#exploration’>below</a>. These variables included age of first incarceration, marital status, personality, highest level of education, and even favorite ice cream flavor. Out of this exploration we coined the term and engineered the Longitudinal Chaos Index (LCI).

=LCI=

The LCI is a measure of the objective chaos of a subject’s love life throughout the study. The more changes in your relationship/marital status, the higher your LCI. To create the LCI feature, we looked at reported marital status over the span of the study. The specific interview question used in the study was “Respondent's marital status in this month in [1994-2016]” (which will now be referred to as the “research question”) and was recorded every month after the subject’s fourteenth birthday. There were six possible responses: “Never Married, Not Cohabitating,” “Never Married, Cohabitating,” “Married,” “Legally Separated,” “Divorced,” and “Widowed.”

The LCI was designed to start at zero for each subject and then increase or decrease according to any change in response to the research question. All but two changes result in an increase by one point. The first exception is the change from “Never Married, Cohabitating” to “Married,” which we determined would not increase a subject’s chaos and therefore does not change the LCI. The second exception is the change from “Married” to “Divorced,” which increases the LCI by two points. After the survey subjects’ LCIs were calculated, we created scoring divisions. These scores detail a final, qualitative measure of a person's LCI. The categories are: “Sad and Alone” (SA), “Happily Married” (HM), “It’s About the Journey” (AJ), “Train Wreck” (TW).

=MACHINE LEARNING=

After defining the LCI, we created a predictive model to predict LCI based on life events. To find the variables that would be the most predictive, we created a simple model that evaluated a single variable. The model was then iterated to loop through each interview question in the longitudinal study to determine the interview questions that most affected LCI. Click here to see the interview questions and our commentary. Then, continue below to calculate your LCI based on your life events, using our predictive model.

|  |  |  |  |
| --- | --- | --- | --- |
| 9 | R3500900  Before 18 | 1999 Have you had sexual intercourse since the last interview on [date of last  interview], that is, made love, had sex, or gone all the way with a person of  the opposite sex? | Yes  No |
| 8 | E7013407  Age 33 (or current) | 2014 MARITAL: MARITAL STATUS IN MONTH 07  Respondent's marital status in this month in 2014 | “Never Married, Not Cohabitating”  “Never Married, Cohabiting”  “Married”  “Legally Separated”  “Divorced”  “Widowed” |
| 7 | S1542000  Age 21 | 2002 RS RELATIONSHIP TO HOUSEHOLD PARENT FIGURE  Relationship of the parent figure(s)/guardian(s) in household to the youth as of  the survey date. | Both biological parents  Two parents, biological mother  Two parents, biological father  Biological mother only  Biological father only  Adoptive parent(s)  Foster parent(s)  No parents, grandparents  No parents, other relatives  Anything else |
| 7 | E8013403  Monthly average at Age 33 (or current) | 2014 ARRESTS: ARREST STATUS IN MONTH 03 | “R not arrested in this month and not arrested in a previous month”  “Number of times R arrested in this month”  “R arrested previously but not in this month” |
| 6 | R1487700  Rank grade school delinquency | DELINQUENCY SCORE INDEX  Youth Report, Delinquency Score Index. Scores range from 0 to 10; higher scores  indicate more incidents of delinquency. | 0  1 TO 2  3 TO 4  5 TO 6  7 TO 8  9 TO 10 |
| 5 | T6670800  At age 30 | TIMING FROM ROUND 15 INTERVIEW  \* Total respondent timing from round 15 interview (measured in seconds). This is  the total interview time excluding the interviewer remarks, the locator section,  and the locator questions in the household information section. | 0  .1 to 1000.0  1000.1 to 2000.0  2000.1 to 3000.0  3000.1 to 4000.0  4000.1 to 5000.0  5000.1 to 6000.0  6000.1 to 7000.0  7000.1 to 8000.0  8000.1 to 9000.0  9000.1 to 10000.0  10000.1 to 20000.0  20000.1 to 30000.0 |
| 5 | R2602100  17  (rank scale 0 to 3)  Age 17 | SUBSTANCE USE INDEX (DLI)  Youth Report, Substance Use Index. Scores range from 0 to 3; higher scores  indicate more instances of substance use since the date of the last interview. 1998 | 0  1  2  3 |
| 2 | S3594800  At age 22 | 2003 Collapsed distance in miles between the respondent's reported address and  the father's reported address. | Lives in the same household  1 to 5 Miles  6 to 10 Miles  11 to 30 Miles  31 to 60 Miles  61 to 100 Miles  101 to 200 Miles  201 to 400 Miles  401 to 700 Miles  >700 Miles |
| 2 | T6216501 | where 1 means disagree strongly and  7 means agree strongly, how much do you agree or disagree that the following  statements describe who you are and how you act.  RESPONSE CHOICE: "When I was in school, I used to break rules quite regularly." | Disagree strongly  Disagree moderately  Disagree a little  Neither agree nor disagree  Agree a little  Agree moderately  Agree strongly |
| 2 | S0921200  Monthly average at age 21 | 2002 HOW OFTEN R DEPRESSED IN LAST MONTH  How much of the time during the last month have you felt so down in the dumps  that nothing could cheer you up? | All of the time  Most of the time  Some of the time  None of the time |
| 2 | R0070700  At age 16 | 1997  PERCENT OF PEERS BELONG TO A GANG that does illegal activities | Almost none (less than 10%)  About 25%  About half (50%)  About 75%  Almost all (more than 90%) |