|  |  |  |  |
| --- | --- | --- | --- |
| **New Variables for DSO562 Project 3**  **(Credit Cards Transactions Data)**  Note: Variables with Red Color are new-created variables. | | | |
| **Description of Variables** | **# Variables Created** | **Approximate CPU Time (Minute)** | **Names of New Columns** |
| Day of Week of Each Application (including Monday to Sunday, 7 unique values) | 1 | <0.1 min | dow |
| Target Encoding for dow column (transferring categorical variables to numerical variables) | 1 | <0.1min | Dow\_risk |
| The categories of amount spent each transaction (6 categories, according to the daily experience) | 1 | <0.1m | Amount\_type |
| Target Encoding for amount type (transferring categorical variables to numerical variables) | 1 | <0.1m | atype\_risk |
| Combination of card number and merchant number, merchant state, merchant zipcode and amount type of each transaction | 4 | <0.1m | Cardnum\_Merchnum  Cardnum\_Merstate  Cardnum\_Merzip  Cardnum\_Atype |
| Combination of amount type and merchant number, merchant state and merchant zipcode for each transaction | 3 | <0.1m | Merchant\_Atype  State\_Atype  Zip\_Atype |
| Number of days since transaction with that entity has been seen. Entities are {Card number, merchant number, Cardnum\_Merchnum, Cardnum\_Merstate, Cardnum\_Merzip, Cardnum\_Atype, Merchant\_Atype, State\_Atype, Zip\_Atype} | 9 | About 20 s | {Cardnum, Merchnum, Cardnum\_Merchnum  Cardnum\_Merstate  Cardnum\_Merzip  Cardnum\_Atype, Merchant\_Atype  State\_Atype  Zip\_Atype }\_day\_since |
| Number of transactions at that entity over the past n days. Entities are {Card number, merchant number, Cardnum\_Merchnum, Cardnum\_Merstate, Cardnum\_Merzip, Cardnum\_Atype, Merchant\_Atype, State\_Atype, Zip\_Atype}, n is {0,1,3,7,14,30} | 54 | About 1 min | {Cardnum, Merchnum, Cardnum\_Merchnum  Cardnum\_Merstate  Cardnum\_Merzip  Cardnum\_Atype, Merchant\_Atype  State\_Atype  Zip\_Atype}\_count\_{0,1,3,7,14,30} |
| {Average, max, median, total, actual/average, actual/max, actual/median, actual/total} transaction amount at that entity over the past n days. Entities are {Card number, merchant number, Cardnum\_Merchnum, Cardnum\_Merstate, Cardnum\_Merzip, Cardnum\_Atype, Merchant\_Atype, State\_Atype, Zip\_Atype}, n is {0,1,3,7,14,30} | 432 | About 3 min | {Cardnum, Merchnum, Cardnum\_Merchnum  Cardnum\_Merstate  Cardnum\_Merzip  Cardnum\_Atype, Merchant\_Atype  State\_Atype  Zip\_Atype}\_{avg, max, med, total, actual/avg, actual/max, axtual/med, actual/toal}\_{0,1,3,7,14,30} |
| Maximum number of transactions under same {Card number, merchant number, Cardnum\_Merchnum, Cardnum\_Merstate, Cardnum\_Merzip, Cardnum\_Atype, Merchant\_Atype, State\_Atype, Zip\_Atype} in the past {3,7,30} days | 27 | About 0.5 min | Max\_count\_by\_{Cardnum, Merchnum, Cardnum\_Merchnum  Cardnum\_Merstate  Cardnum\_Merzip  Cardnum\_Atype, Merchant\_Atype  State\_Atype  Zip\_Atype}\_{3,7,30} |
| Number of transactions at that entity seen in a short time window {today or past day} compared to number of transactions at that entity over a longer time window {past 3, 7,14,30 days}. Entities are {Card number, merchant number, Cardnum\_Merchnum, Cardnum\_Merstate, Cardnum\_Merzip, Cardnum\_Atype, Merchant\_Atype, State\_Atype, Zip\_Atype} | 72 | About 34 s | {Cardnum, Merchnum, Cardnum\_Merchnum  Cardnum\_Merstate  Cardnum\_Merzip  Cardnum\_Atype, Merchant\_Atype  State\_Atype  Zip\_Atype}\_count\_{0,1}\_by\_{3,7,14,30} |
| Number of unique identy in the specific field during the past {1,3,7,14,30,60} days. Both entity and field are {Card number, merchant number, Cardnum\_Merchnum, Cardnum\_Merstate, Cardnum\_Merzip, Cardnum\_Atype, Merchant\_Atype, State\_Atype, Zip\_Atype}, but with different values. | 432 | About 32 mins | { Cardnum, Merchnum, Cardnum\_Merchnum  Cardnum\_Merstate  Cardnum\_Merzip  Cardnum\_Atype, Merchant\_Atype  State\_Atype  Zip\_Atype}\_for\_{1,3,7,14,30,60} |
| Benford Law’s variables for each card number and each merchant number except for Fedex transactions | 2 | About 40 s | Cardnum\_b,  Merchnum\_b |
| Total | 1039 | About 40 mins |  |