**As of 11/8**

Small update – demographics includes personid, 37 attributes for race, 3 for gender, 3 for ethnicity, one for age, one for deceased and the longCovid attribute.

Biggest update to the pipeline – created new datasets just last week – still running (but the code works)

|  |  |
| --- | --- |
| # of Long Covid Identified Concept Ids @ | ConceptIds |
| ICHI (2) | 705076, 710706 |
| Fall 2023 (9) | 600715,  710708,  705076,  608525,  766503,  766502,  600589,  710706,  600588 |

Condtionoccurrence – (27 selected)

["conditionduration, age, 37311061, 257011, 444094, 4188598, 0, 437663, 434613, 140673, 195867, 4193704, 4244438, 439658, 4147829, 196523, 4042889, 4239938, 43530950, 197607, 601619, 140214, 436962, 436665, 435796, 141825, 4138617]

Conditionera- (31 selected)

[age, 37311061, 4273307, 254761, 257011, 25297, 320128, 442077, 378253, 77670, 433736, 200219, 437663, 4223659, 440029, 434005, 436096, 194133, 45768910, 4188598, 440383, 4081648, 434169, 432545, 194696, 601619, 37018196, 257012, 435655, 134898, 705076]

Procedures – (31 selected)

[2313814, 2514408, 40756910, 2211767, 2514526, 42628602, 4036803, 46257484, 2211488, 2213221, 2211747, 9202, 2211749, 2514436, 2514435, 2108887, 2211737, 4141760, 42627940, 43527985, 42627952, 42742498, 2211826, "age, 2514527, 725069, 2313816, 2414397, 705050, 2414398, 46257421]

Devices – (27 selected)

["age, 2004208005, 2004208004, 2615309, 40664904, 2615762, 2615740, 4224038, 45110833, 2720522, 4159813, 2614958, 44786360, 2719876, 4145694, 2720868, 4097216, 2720869, 4145528, 2614897, 2615362, 2615358, 2614751, 2616482, 2614959, 2615324, 2719755]

Drugs - (30 selected)

[drug\_era\_duration, age, 1759206, 19135374, 19079524, 19080985, 1154343, 1127433, 19049105, 1734104, 46234346, 1326115, 1124957, 19036781, 40227012, 19127890, 40048832, 40232756, 1759842, 1521369, 753626, 1129625, 732893, 778711, 1589505, 19005968, 43012671, 42800267, 791967, 1717704]

Takes the longest

Observation- (32 selected)

[age, 37311059, 1340204, 3040464, 37016200, 4283657, 4110772, 46272450, 443364, 36307579, 36307819, 4041306, 4188893, 3018063, 36305168, 4256640, 4224504, 40217302, 45881517, 40481872, 4214956, 36304041, 4167217, 4060705, 441482, 46273937, 45765650, 45773394, 4005823, 2004207791, 3027028, 927173]

Measurement - ()

Conditions - (58 selected)  
["conditionduration, age, 37311061, 257011, 444094, 4188598, 0, 437663, 434613, 140673, 195867, 4193704, 4244438, 439658, 4147829, 196523, 4042889, 4239938, 43530950, 197607, 601619, 140214, 436962, 436665, 435796, 141825, 4138617, 37311061, 4273307, 254761, 257011, 25297, 320128, 442077, 378253, 77670, 433736, 200219, 437663, 4223659, 440029, 434005, 436096, 194133, 45768910, 4188598, 440383, 4081648, 434169, 432545, 194696, 601619, 37018196, 257012, 435655, 134898, 705076]

Having memory issues since the vectorized features are huge but again code works  
  
LightGBM Models – devices & procedures results (LightGBMallFeats sheet)

[https://txst-my.sharepoint.com/:x:/r/personal/jt463txstateedu/layouts/15/Doc.aspx?sourcedoc=%7BFF9BAEFC-4028-4BB0-8ABF-037D99C4280E%7D&file=dataSumFeatSelectModScore1123.xlsx&action=default&mobileredirect=true](https://txst-my.sharepoint.com/:x:/r/personal/j_t463_txstate_edu/_layouts/15/Doc.aspx?sourcedoc=%7BFF9BAEFC-4028-4BB0-8ABF-037D99C4280E%7D&file=dataSum_FeatSelect_ModScore_1123.xlsx&action=default&mobileredirect=true)