![](data:application/octet-stream;base64,)

CDM Onboarding report for the Synthea synthetic health database database

Maxim Moinat

1. Table of content
2. Execution details

| Detail | Value |
| --- | --- |
| CdmOnboarding package Version | 1.0.0.900 |
| CDM version | v5.3 |
| Execution date | Tue Jul 19 16:23:55 2022 |
| Execution duration | 67.40 seconds |
| Document Generation date | Tue Jul 19 16:23:55 2022 |

* 1. CDM Source Table

1. Content of the OMOP cdm\_source table

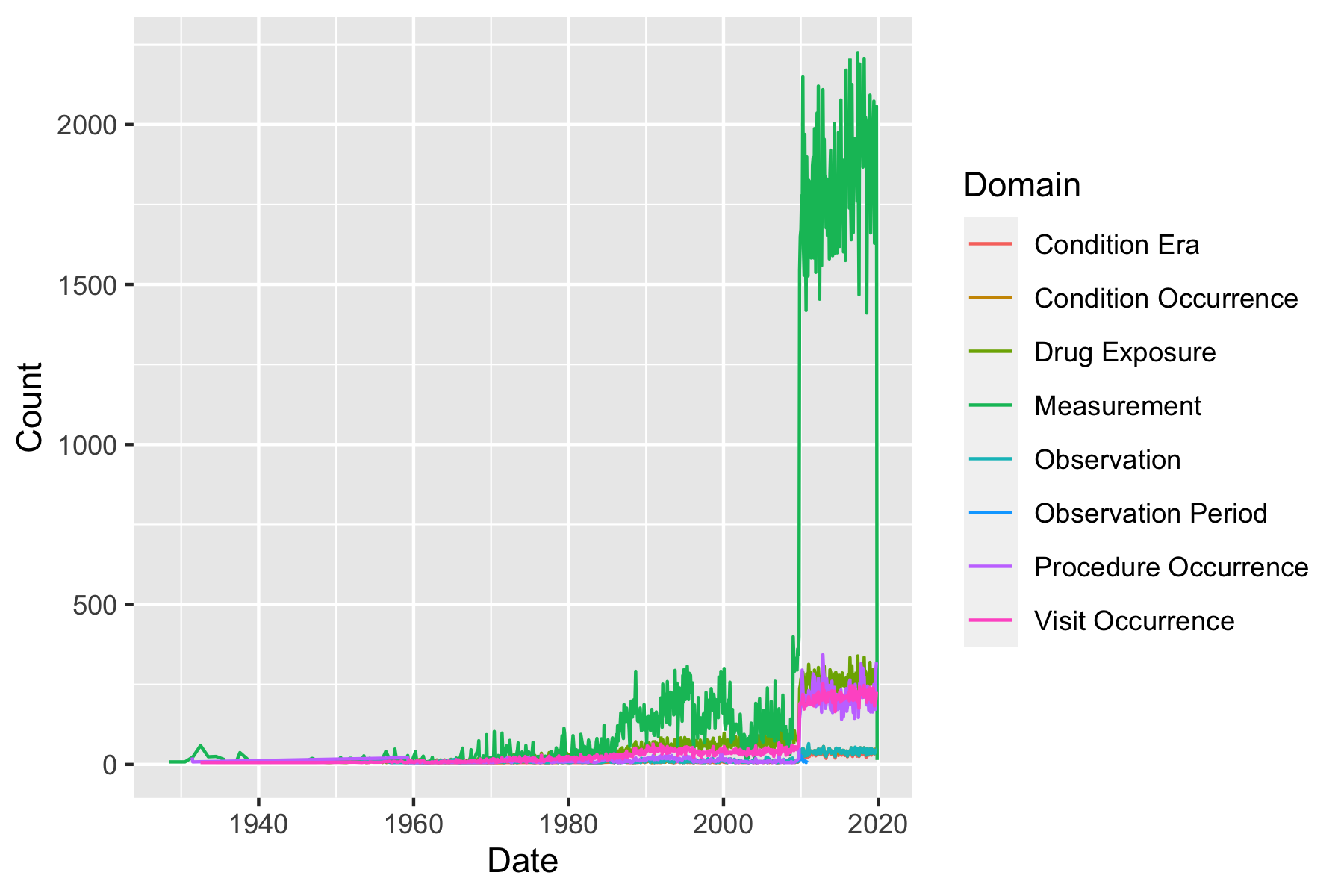
| field | 1 |
| --- | --- |
| CDM\_SOURCE\_NAME | Synthea synthetic health database |
| CDM\_SOURCE\_ABBREVIATION | Synthea |
| CDM\_HOLDER | OHDSI Community |
| SOURCE\_DESCRIPTION | SyntheaTM is a Synthetic Patient Population Simulator. The goal is to output synthetic, realistic (but not real), patient data and associated health records in a variety of formats. |
| SOURCE\_DOCUMENTATION\_REFERENCE | https://synthetichealth.github.io/synthea/ |
| CDM\_ETL\_REFERENCE | https://github.com/OHDSI/ETL-Synthea |
| SOURCE\_RELEASE\_DATE | 2021-10-04 |
| CDM\_RELEASE\_DATE | 2021-10-04 |
| CDM\_VERSION | v5.3 |
| VOCABULARY\_VERSION | v5.0 21-DEC-20 |

1. Clinical data
   1. Record counts per OMOP CDM table
2. The number of records in all clinical data tables

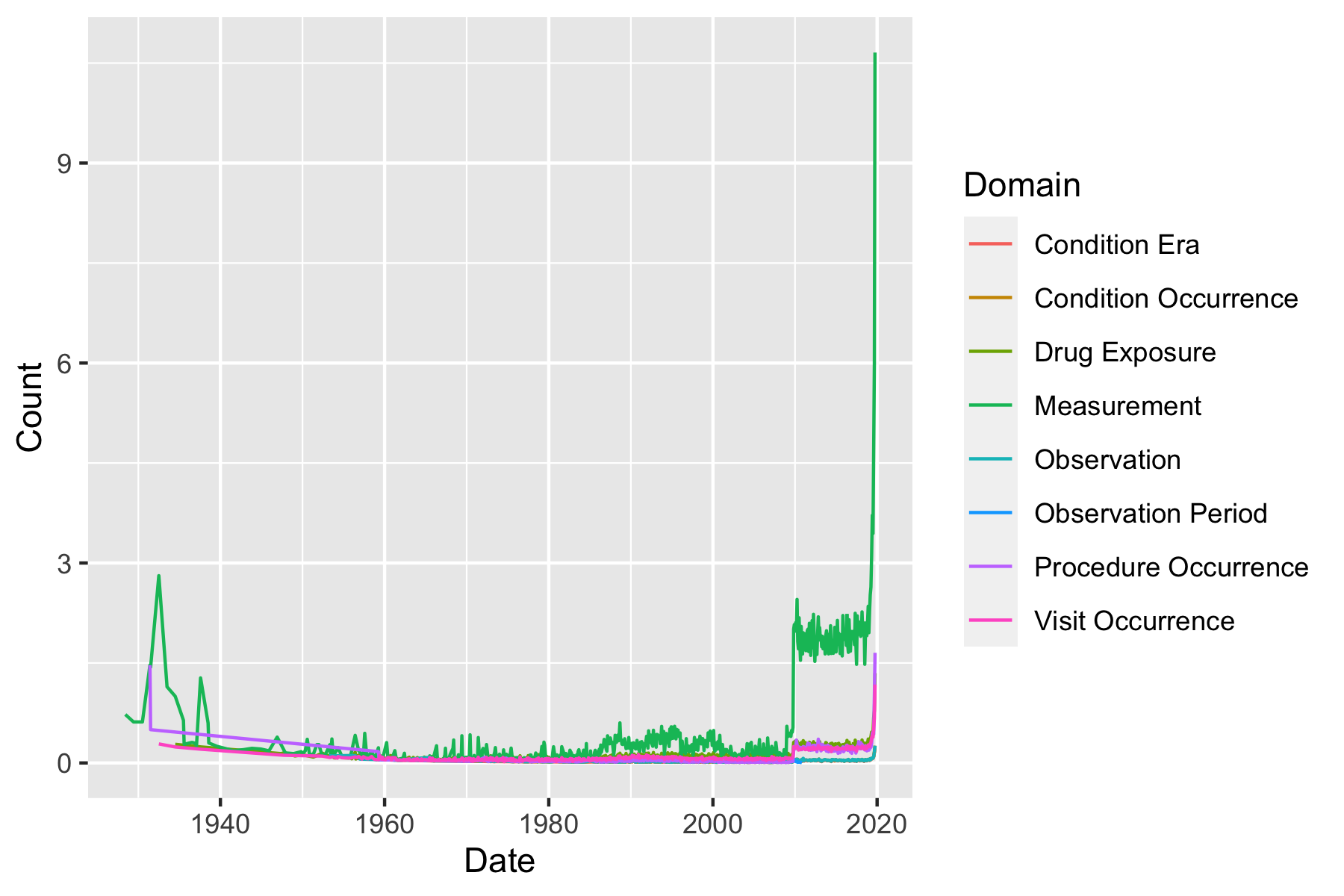
| TABLENAME | COUNT | PERSON COUNT |
| --- | --- | --- |
| measurement | 401,655 | 2,778 |
| visit\_detail | 261,710 | 25,845 |
| visit\_occurrence | 261,710 | 25,845 |
| drug\_exposure | 129,217 | 4,643 |
| person | 25,845 | 25,845 |
| observation\_period | 25,845 | 25,845 |
| drug\_era | 17,856 | 4,643 |
| procedure\_occurrence | 15,568 | 2,117 |
| condition\_era | 8,170 | 5,335 |
| condition\_occurrence | 8,170 | 5,335 |
| observation | 2,461 | 1,786 |
| device\_exposure | 1,480 | 1,445 |
| death | 845 | 845 |
| location | 0 | NA |
| cost | 0 | NA |
| care\_site | 0 | NA |
| note | 0 | 0 |
| payer\_plan\_period | 0 | 0 |
| provider | 0 | NA |
| specimen | 0 | 0 |
| dose\_era | 0 | 0 |

Query executed in 0.95 seconds

* 1. Data density plots



1. Total record count over time per OMOP data domain



1. Number of records per person over time per OMOP data domain
   1. Distinct concepts per person
2. The number of distinct concepts per person for all OMOP data domains

| Domain | Min | P10 | P25 | MEDIAN | P75 | P90 | Max |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Visit Occurrence | 1 | 1 | 1 | 2 | 2 | 3 | 3 |
| Condition Occurrence | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Procedure Occurrence | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Drug exposure | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Observation | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Condition era | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Measurement | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

1. Vocabulary data
   1. Vocabularies

Vocabulary version: v5.0 21-DEC-20

1. The vocabularies available in the CDM with concept count. Note that this does not reflect which concepts are actually used in the clinical CDM tables. S=Standard, C=Classification and '-'=Non-standard

| ID | NAME | VERSION | S | C | - |
| --- | --- | --- | --- | --- | --- |
| ABMS | Provider Specialty (American Board of Medical Specialties) | 2018-06-26 ABMS | 87 | 0 | 11 |
| ATC | WHO Anatomic Therapeutic Chemical Classification | RxNorm 20200504 | 0 | 6,335 | 61 |
| CANCERTYPE | DIAMOND cancer type for family history | v0.1 | 0 | 0 | 1 |
| CCAM | Common Classification of Medical Acts | CCAM version 64 | 0 | 0 | 10,206 |
| CDM | OMOP Common DataModel | CDM v6.0.2 | 950 | 0 | 0 |
| CMS Place of Service | Place of Service Codes for Professional Claims (CMS) | 2009-01-11 | 51 | 0 | 9 |
| COMORBIDITIES | DIAMOND comorbidities | v0.1 | 0 | 0 | 1 |
| CPT4 | Current Procedural Terminology version 4 (AMA) | 2020 Release | 12,686 | 3,433 | 173 |
| Cancer Modifier | Diagnostic Modifiers of Cancer (OMOP) | Cancer Modifier 20201014 | 3,304 | 0 | 33 |
| Cohort Type | OMOP Cohort Type | NA | 0 | 0 | 1 |
| Concept Class | OMOP Concept Class | NA | 0 | 0 | 414 |
| Condition Status | OMOP Condition Status | NA | 22 | 0 | 0 |
| Condition Type | OMOP Condition Occurrence Type | NA | 0 | 0 | 118 |
| Cost | OMOP Cost | NA | 51 | 0 | 0 |
| Cost Type | OMOP Cost Type | NA | 0 | 0 | 8 |
| Currency | International Currency Symbol (ISO 4217) | 2008 | 180 | 0 | 0 |
| Death Type | OMOP Death Type | NA | 0 | 0 | 14 |
| Device Type | OMOP Device Type | NA | 0 | 0 | 4 |
| Domain | OMOP Domain | NA | 0 | 0 | 64 |
| Drug Type | OMOP Drug Exposure Type | NA | 0 | 0 | 16 |
| ETHNICITY | DIAMOND ethnicity | v0.1 | 0 | 0 | 1 |
| Episode | OMOP Episode | Episode 20201014 | 14 | 0 | 4 |
| Ethnicity | OMOP Ethnicity | NA | 2 | 0 | 0 |
| FAMILY\_RELATIONSHIP | DIAMOND family relationship for family history | v0.1 | 0 | 0 | 1 |
| Gender | OMOP Gender | NA | 2 | 0 | 3 |
| HCPCS | Healthcare Common Procedure Coding System (CMS) | 2020 Alpha Numeric HCPCS File | 8,434 | 0 | 2,355 |
| HES Specialty | Hospital Episode Statistics Specialty (NHS) | 2018-06-26 HES Specialty | 64 | 0 | 101 |
| HGNC | Human Gene Nomenclature (European Bioinformatics Institute) | HGNC v20200101 | 19,297 | 0 | 0 |
| HemOnc | HemOnc | HemOnc 2019-08-29 | 1,730 | 343 | 4,962 |
| ICD10 | International Classification of Diseases, Tenth Revision (WHO) | 2020 Release | 0 | 0 | 16,143 |
| ICD10CM | International Classification of Diseases, Tenth Revision, Clinical Modification (NCHS) | ICD10CM FY2020 code descriptions | 0 | 0 | 96,308 |
| ICD9CM | International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 1 and 2 (NCHS) | ICD9CM v32 master descriptions | 0 | 0 | 17,564 |
| ICD9Proc | International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 3 (NCHS) | ICD9CM v32 master descriptions | 2,237 | 0 | 2,420 |
| ICDO3 | International Classification of Diseases for Oncology, Third Edition (WHO) | ICDO3 SEER Site/Histology Released 06/2019 | 59,619 | 0 | 4,048 |
| IMAGING | DIAMOND Imaging | v0.1 | 0 | 0 | 1 |
| JAX | The Clinical Knowledgebase (The Jackson Laboratory) | JAX v20200824 | 0 | 0 | 5,877 |
| Korean Revenue Code | Korean Revenue Code | NA | 7 | 0 | 0 |
| LOINC | Logical Observation Identifiers Names and Codes (Regenstrief Institute) | 2.68 | 105,939 | 46,354 | 97,994 |
| Meas Type | OMOP Measurement Type | NA | 0 | 0 | 12 |
| Medicare Specialty | Medicare provider/supplier specialty codes (CMS) | 2018-06-26 Specialty | 107 | 0 | 13 |
| Metadata | Metadata | NA | 1 | 0 | 1 |
| NAACCR | Data Standards & Data Dictionary Volume II (NAACCR) | NAACCR v18 | 26,240 | 0 | 8,233 |
| NCIt | NCI Thesaurus (National Cancer Institute) | NCIt 2018-01-01 | 1,899 | 0 | 527 |
| NDC | National Drug Code (FDA and manufacturers) | NDC 20201129 | 10,617 | 0 | 1,012,135 |
| NUCC | National Uniform Claim Committee Health Care Provider Taxonomy Code Set (NUCC) | 2018-06-26 NUCC | 697 | 0 | 158 |
| None | OMOP Standardized Vocabularies | v5.0 21-DEC-20 | 0 | 0 | 1 |
| Note Type | OMOP Note Type | NA | 0 | 0 | 10 |
| OMOP Extension | OMOP Extension (OHDSI) | OMOP Extension 20201217 | 67,541 | 0 | 1 |
| OPCS4 | OPCS Classification of Interventions and Procedures version 4 (NHS) | 2019 Release | 2,372 | 0 | 8,628 |
| OSM | OpenStreetMap | OSM Release 2019-02-21 | 203,339 | 0 | 0 |
| Obs Period Type | OMOP Observation Period Type | NA | 0 | 0 | 6 |
| Observation Type | OMOP Observation Type | NA | 0 | 0 | 29 |
| PAST\_PROCEDURE | DIAMOND past procedures | v0.1 | 0 | 0 | 1 |
| PHDSC | Source of Payment Typology (PHDSC) | Version 3.0 | 158 | 0 | 4 |
| PIONEER | European Network of Excellence for Big Data in Prostate Cancer | v0.1 | 0 | 0 | 1 |
| Plan | Health Plan - contract to administer healthcare transactions by the payer, facilitated by the sponsor | NA | 11 | 0 | 0 |
| Plan Stop Reason | Plan Stop Reason - Reason for termination of the Health Plan | NA | 13 | 0 | 0 |
| Procedure Type | OMOP Procedure Occurrence Type | NA | 0 | 0 | 97 |
| Race | Race and Ethnicity Code Set (USBC) | Version 1.0 | 50 | 0 | 3 |
| Read | NHS UK Read Codes Version 2 (HSCIC) | NHS READV2 21.0.0 20160401000001 + DATAMIGRATION\_25.0.0\_20180403000001 | 0 | 0 | 108,696 |
| Relationship | OMOP Relationship | NA | 14 | 0 | 608 |
| Revenue Code | UB04/CMS1450 Revenue Codes (CMS) | 2010 Release | 538 | 0 | 0 |
| RxNorm | RxNorm (NLM) | RxNorm 20200803 | 147,283 | 36,007 | 109,381 |
| RxNorm Extension | RxNorm Extension (OHDSI) | RxNorm Extension 2020-12-10 | 1,761,116 | 0 | 274,381 |
| SNOMED | Systematic Nomenclature of Medicine - Clinical Terms (IHTSDO) | SNOMED 2020-04-01 | 540,077 | 0 | 346,408 |
| SPL | Structured Product Labeling (FDA) | NDC 20201129 | 0 | 474,244 | 14,893 |
| Sponsor | Sponsor - institution or individual financing healthcare transactions | NA | 6 | 0 | 0 |
| TNM | DIAMOND TNM staging | v0.1 | 0 | 0 | 1 |
| TREATMENT | DIAMOND treatment | v0.1 | 0 | 0 | 1 |
| Type Concept | OMOP Type Concept | NA | 78 | 0 | 0 |
| UB04 Point of Origin | UB04 Claim Source Inpatient Admission Code (CMS) | NA | 0 | 0 | 22 |
| UB04 Pri Typ of Adm | UB04 Claim Inpatient Admission Type Code (CMS) | NA | 6 | 0 | 0 |
| UB04 Pt dis status | UB04 Patient Discharge Status Code (CMS) | NA | 0 | 0 | 55 |
| UB04 Typ bill | UB04 Type of Bill - Institutional (USHIK) | NA | 4 | 0 | 294 |
| UCUM | Unified Code for Units of Measure (Regenstrief Institute) | Version 1.8.2 | 918 | 0 | 86 |
| UK Biobank | UK Biobank | version 2020-10-15 | 4,787 | 291 | 13,359 |
| US Census | United States Census Bureau | US Census 2017 Release | 13 | 0 | 0 |
| Visit | OMOP Visit | NA | 16 | 0 | 0 |
| Visit Type | OMOP Visit Type | NA | 0 | 0 | 18 |
| Vocabulary | OMOP Vocabulary | NA | 0 | 0 | 137 |
| dm+d | Dictionary of Medicines and Devices (NHS) | dm+d Version 12.0.0 20201207 | 22,927 | 0 | 358,911 |

Query executed in 7.47 seconds

* 1. Table counts

1. Shows the number of records in all vocabulary tables

| TABLENAME | COUNT |
| --- | --- |
| concept\_ancestor | 69,595,970 |
| concept\_relationship | 44,604,476 |
| concept\_synonym | 7,573,232 |
| concept | 6,088,528 |
| concept\_class | 416 |
| vocabulary | 81 |
| domain | 48 |
| drug\_strength | 0 |

Query executed in 17.63 seconds

* 1. Mapping Completeness

1. Shows the percentage of codes that are mapped to the standardized vocabularies as well as the percentage of records.

| DOMAIN | #CODES SOURCE | #CODES MAPPED | %CODES MAPPED | #RECORDS SOURCE | #RECORDS MAPPED | %RECORDS MAPPED |
| --- | --- | --- | --- | --- | --- | --- |
| Condition | 19 | 19 | 100 | 8,170 | 8,170 | 100.00 |
| Condition status | 1 | 0 | 0 | 8,170 | 0 | 0.00 |
| Death cause | 6 | 6 | 100 | 845 | 845 | 100.00 |
| Device | 3 | 3 | 100 | 1,480 | 1,480 | 100.00 |
| Drug | 26 | 26 | 100 | 129,217 | 129,217 | 100.00 |
| Measurement | 53 | 53 | 100 | 401,655 | 401,655 | 100.00 |
| Measurement unit | 15 | 12 | 80 | 401,655 | 336,489 | 83.78 |
| Measurement value | 3,560 | 0 | 0 | 401,655 | 0 | 0.00 |
| Observation | 5 | 5 | 100 | 2,461 | 2,461 | 100.00 |
| Observation unit | 1 | 0 | 0 | 2,461 | 0 | 0.00 |
| Observation value | 1 | 0 | 0 | 2,461 | 0 | 0.00 |
| Procedure | 21 | 21 | 100 | 15,568 | 15,568 | 100.00 |
| Provider Specialty | 0 | NA | NA | NA | NA | NA |
| Specimen | 0 | NA | NA | NA | NA | NA |
| Visit | 261,710 | 261,710 | 100 | 261,710 | 261,710 | 100.00 |

Query executed in 2.24 seconds

* 1. Drug Mappings

1. The level of the drug mappings

| CLASS | #RECORDS | #PATIENTS | #SOURCE CODES |
| --- | --- | --- | --- |
| Clinical Drug | 118,270 | 4,321 | 19 |
| Clinical Drug Comp | 9,078 | 664 | 1 |
| Quant Clinical Drug | 1,869 | 779 | 6 |

Query executed in 1.02 seconds

* 1. Unmapped Codes

1. Omitted because no unmapped drugs were found

Query executed in

1. Omitted because no unmapped conditions were found

Query executed in

1. Omitted because no unmapped measurements were found

Query executed in

1. Omitted because no unmapped observations were found

Query executed in

1. Omitted because no unmapped procedures were found

Query executed in

1. Omitted because no unmapped devices were found

Query executed in

1. Omitted because no unmapped visits were found

Query executed in

1. Omitted because no unmapped units were found

Query executed in

* 1. Mapped Codes

1. Top 25 of mapped drugs. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | simvastatin 10 MG Oral Tablet | 37,000 | 2,500 |
| 2 | nitroglycerin 0.4 MG/ACTUAT Mucosal Spray | 14,800 | 1,300 |
| 3 | simvastatin 20 MG Oral Tablet | 13,300 | 1,100 |
| 4 | amlodipine 5 MG Oral Tablet | 12,700 | 1,100 |
| 5 | clopidogrel 75 MG Oral Tablet | 12,100 | 1,600 |
| 6 | warfarin sodium 5 MG Oral Tablet | 9,400 | 700 |
| 7 | digoxin 0.125 MG Oral Tablet | 9,400 | 700 |
| 8 | verapamil hydrochloride 40 MG | 9,100 | 700 |
| 9 | cisplatin 50 MG Injection | 4,300 | 100 |
| 10 | paclitaxel 100 MG Injection | 3,500 | 100 |
| 11 | etoposide 100 MG Injection | 800 | 100 |
| 12 | alteplase 100 MG Injection | 600 | 600 |
| 13 | 1 ML docetaxel 20 MG/ML Injection | 400 | 400 |
| 14 | 0.25 ML leuprolide acetate 30 MG/ML Prefilled Syringe | 400 | 400 |
| 15 | acetaminophen 750 MG / hydrocodone bitartrate 7.5 MG Oral Tablet | 400 | 400 |
| 16 | 1 ML atropine sulfate 1 MG/ML Injection | 400 | 400 |
| 17 | 1 ML epinephrine 1 MG/ML Injection | 400 | 400 |
| 18 | 3 ML amiodarone hydrochloride 50 MG/ML Prefilled Syringe | 300 | 300 |
| 19 | atorvastatin 80 MG Oral Tablet | 300 | 300 |
| 20 | 24 HR nicotine 0.292 MG/HR Transdermal System | 200 | 200 |
| 21 | fluoxetine 20 MG Oral Capsule | 200 | 200 |
| 22 | sertraline 100 MG Oral Tablet | 100 | 100 |
| 23 | acetaminophen 300 MG / hydrocodone bitartrate 10 MG Oral Tablet | 100 | 100 |
| 24 | acetaminophen 325 MG / oxycodone hydrochloride 5 MG Oral Tablet | 100 | 100 |
| 25 | meperidine hydrochloride 50 MG Oral Tablet | 100 | 100 |

Query executed in

1. All 18 mapped conditions. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Hyperlipidemia | 2,500 | 2,500 |
| 2 | Coronary arteriosclerosis | 1,300 | 1,300 |
| 3 | Cerebrovascular accident | 1,100 | 1,100 |
| 4 | Cardiac arrest | 900 | 900 |
| 5 | Atrial fibrillation | 700 | 700 |
| 6 | Myocardial infarction | 500 | 500 |
| 7 | Malignant tumor of prostate | 500 | 500 |
| 8 | Carcinoma in situ of prostate | 400 | 400 |
| 9 | Major depressive disorder | 200 | 200 |
| 10 | Alcoholism | 200 | 200 |
| 11 | Opioid abuse | 200 | 200 |
| 12 | Metastasis from malignant tumor of prostate | 100 | 100 |
| 13 | Non-small cell carcinoma of lung, TNM stage 1 | 100 | 100 |
| 14 | Non-small cell lung cancer | 100 | 100 |
| 15 | Major depression, single episode | 100 | 100 |
| 16 | Posttraumatic stress disorder | 100 | 100 |
| 17 | Primary small cell malignant neoplasm of lung, TNM stage 1 | 100 | 100 |
| 18 | Small cell carcinoma of lung | 100 | 100 |

Query executed in

1. Top 25 of mapped measurements. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Calcium [Mass/volume] in Blood | 19,500 | 2,600 |
| 2 | Sodium [Moles/volume] in Blood | 19,500 | 2,600 |
| 3 | Alkaline phosphatase [Enzymatic activity/volume] in Serum or Plasma | 19,500 | 2,600 |
| 4 | Potassium [Moles/volume] in Blood | 19,500 | 2,600 |
| 5 | Protein [Mass/volume] in Serum or Plasma | 19,500 | 2,600 |
| 6 | Carbon dioxide, total [Moles/volume] in Blood | 19,500 | 2,600 |
| 7 | Aspartate aminotransferase [Enzymatic activity/volume] in Serum or Plasma | 19,500 | 2,600 |
| 8 | Creatinine [Mass/volume] in Blood | 19,500 | 2,600 |
| 9 | Bilirubin.total [Mass/volume] in Serum or Plasma | 19,500 | 2,600 |
| 10 | Albumin [Mass/volume] in Serum or Plasma | 19,500 | 2,600 |
| 11 | Chloride [Moles/volume] in Blood | 19,500 | 2,600 |
| 12 | Alanine aminotransferase [Enzymatic activity/volume] in Serum or Plasma | 19,500 | 2,600 |
| 13 | Globulin [Mass/volume] in Serum by calculation | 19,500 | 2,600 |
| 14 | Urea nitrogen [Mass/volume] in Blood | 19,500 | 2,600 |
| 15 | Glomerular filtration rate/1.73 sq M.predicted [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (MDRD) | 19,500 | 2,600 |
| 16 | Glucose [Mass/volume] in Blood | 19,500 | 2,600 |
| 17 | Cholesterol in LDL [Mass/volume] in Serum or Plasma by Direct assay | 19,200 | 2,500 |
| 18 | Triglyceride [Mass/volume] in Serum or Plasma | 19,200 | 2,500 |
| 19 | Cholesterol [Mass/volume] in Serum or Plasma | 19,200 | 2,500 |
| 20 | Cholesterol in HDL [Mass/volume] in Serum or Plasma | 19,200 | 2,500 |
| 21 | Physical findings of Prostate | 2,100 | 400 |
| 22 | Prostate specific Ag [Mass/volume] in Serum or Plasma | 2,100 | 400 |
| 23 | Appearance of Urine | 700 | 200 |
| 24 | Mental health Outpatient Note | 500 | 100 |
| 25 | US Guidance for biopsy of Prostate | 400 | 300 |

Query executed in

1. All 5 mapped observations. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | History of cardiac arrest | 900 | 900 |
| 2 | Cause of death [US Standard Certificate of Death] | 900 | 900 |
| 3 | History of myocardial infarction | 500 | 500 |
| 4 | Smokes tobacco daily | 200 | 200 |
| 5 | Suspected lung cancer | 100 | 100 |

Query executed in

1. All 21 mapped procedures. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Combined chemotherapy and radiation therapy | 4,300 | 100 |
| 2 | Direct current cardioversion | 4,000 | 700 |
| 3 | Digital examination of rectum | 2,100 | 400 |
| 4 | Echocardiography | 1,400 | 1,300 |
| 5 | Catheter ablation of tissue of heart | 1,300 | 800 |
| 6 | Percutaneous mechanical thrombectomy of portal vein using fluoroscopic guidance | 600 | 600 |
| 7 | Percutaneous coronary intervention | 500 | 500 |
| 8 | Coronary artery bypass graft | 500 | 500 |
| 9 | Biopsy of prostate | 400 | 300 |
| 10 | Insertion of biventricular implantable cardioverter defibrillator | 400 | 400 |
| 11 | Prostatectomy | 200 | 200 |
| 12 | Magnetic resonance imaging for measurement of brain volume | 100 | 100 |
| 13 | CT of chest and abdomen | 100 | 100 |
| 14 | Plain chest X-ray | 100 | 100 |
| 15 | Depression screening using PHQ-9 (Patient Health Questionnaire 9) score | 100 | 100 |
| 16 | Diagnostic fiberoptic bronchoscopy | 100 | 100 |
| 17 | Suicide risk assessment | 100 | 100 |
| 18 | Fine needle aspiration biopsy of lung | 100 | 100 |
| 19 | Thoracentesis | 100 | 100 |
| 20 | Sputum microscopy | 100 | 100 |
| 21 | Mental health screening | 100 | 100 |

Query executed in

1. All 3 mapped devices. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Implantable defibrillator | 900 | 900 |
| 2 | Coronary artery stent | 500 | 500 |
| 3 | Implantable cardiac pacemaker | 200 | 200 |

Query executed in

1. All 3 mapped visits. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Outpatient Visit | 255,000 | 25,900 |
| 2 | Inpatient Visit | 4,300 | 100 |
| 3 | Emergency Room Visit | 2,600 | 2,300 |

Query executed in

1. Omitted because no mapped units were found

Query executed in

* 1. Source to concept map

1. Source to concept map breakdown

| SOURCE\_VOCABULARY\_ID | TARGET\_VOCABULARY\_ID | COUNT |
| --- | --- | --- |
| CANCERTYPE | None | 38 |
| ETHNICITY | None | 11 |
| FAMILY\_RELATIONSHIP | None | 120 |
| IMAGING | None | 33 |
| TNM | None | 40 |
| TREATMENT | None | 39 |

Query executed in 0.14 seconds

Note that the full source\_to\_concept\_map table is added in the results.zip

1. Technical Infrastructure
   1. HADES packages
2. Versions of all installed R packages that are relevant for DARWIN EU studies

| Package | Version |
| --- | --- |
| Andromeda | 0.6.1 |
| BigKnn | 1.0.1 |
| Capr | 1.0.3 |
| CirceR | 1.1.1 |
| CohortGenerator | 0.6.0 |
| CohortMethod | 4.2.2 |
| Cyclops | 3.2.0 |
| DatabaseConnector | 5.0.4 |
| EmpiricalCalibration | 3.1.0 |
| Eunomia | 1.0.2 |
| EvidenceSynthesis | 0.2.3 |
| FeatureExtraction | 3.2.0 |
| Hydra | 0.3.0 |
| MethodEvaluation | 2.2.0 |
| OhdsiSharing | 0.2.2 |
| ParallelLogger | 3.0.1 |
| PhenotypeLibrary | 3.1.2 |
| ROhdsiWebApi | 1.3.1 |
| SelfControlledCaseSeries | 3.3.0 |
| SelfControlledCohort | 1.6.0 |
| SqlRender | 1.9.2 |

Missing R packages: PatientLevelPrediction, EnsemblePatientLevelPrediction, CohortDiagnostics

* 1. System Information

Installed R version: R version 4.0.3 (2020-10-10)

System CPU vendor:

System CPU model: Apple M1

System CPU number of cores: 8

System RAM: 17.18 GB

DBMS: postgresql

WebAPI version: 2.11.0

* 1. Vocabulary Query Performance

The number of 'Maps To' relations is equal to 4847747, queried in 28.23 seconds (5.82327e-06 s/#).

* 1. Achilles Query Performance

1. Execution time of queries of the Achilles R-Package

| ID | NAME | DURATION |
| --- | --- | --- |
| 0 | Source name | 0.063486 secs |
| 1 | Number of persons | 0.012376 secs |
| 2 | Number of persons by gender | 0.018776 secs |
| 3 | Number of persons by year of birth | 0.020583 secs |
| 4 | Number of persons by race | 0.011756 secs |
| 5 | Number of persons by ethnicity | 0.013231 secs |
| 7 | Number of persons with invalid provider\_id | 0.013621 secs |
| 8 | Number of persons with invalid location\_id | 0.018225 secs |
| 9 | Number of persons with invalid care\_site\_id | 0.012805 secs |
| 10 | Number of all persons by year of birth by gender | 0.019509 secs |
| 11 | Number of non-deceased persons by year of birth by gender | 0.015394 secs |
| 12 | Number of persons by race and ethnicity | 0.013079 secs |
| 101 | Number of persons by age, with age at first observation period | 0.027144 secs |
| 102 | Number of persons by gender by age, with age at first observation period | 0.023618 secs |
| 108 | Number of persons by length of observation period, in 30d increments | 0.020133 secs |
| 109 | Number of persons with continuous observation in each year | 0.086811 secs |
| 110 | Number of persons with continuous observation in each month | 0.203220 secs |
| 111 | Number of persons by observation period start month | 0.017839 secs |
| 112 | Number of persons by observation period end month | 0.010727 secs |
| 113 | Number of persons by number of observation periods | 0.014020 secs |
| 114 | Number of persons with observation period before year-of-birth | 0.012052 secs |
| 115 | Number of persons with observation period end < observation period start | 0.010412 secs |
| 116 | Number of persons with at least one day of observation in each year by gender and age decile | 0.142697 secs |
| 117 | Number of persons with at least one day of observation in each month | 7.148543 secs |
| 118 | Number of observation periods with invalid person\_id | 0.012633 secs |
| 119 | Number of observation period records by period\_type\_concept\_id | 0.012655 secs |
| 200 | Number of persons with at least one visit occurrence, by visit\_concept\_id | 0.124208 secs |
| 201 | Number of visit occurrence records, by visit\_concept\_id | 0.061951 secs |
| 202 | Number of persons by visit occurrence start month, by visit\_concept\_id | 0.194482 secs |
| 204 | Number of persons with at least one visit occurrence, by visit\_concept\_id by calendar year by gender by age decile | 0.194499 secs |
| 207 | Number of visit records with invalid person\_id | 0.027467 secs |
| 208 | Number of visit records outside valid observation period | 0.035518 secs |
| 209 | Number of visit records with end date < start date | 0.017478 secs |
| 210 | Number of visit records with invalid care\_site\_id | 0.018506 secs |
| 212 | Number of persons with at least one visit occurrence, by calendar year by gender by age decile | 0.134329 secs |
| 220 | Number of visit occurrence records by visit occurrence start month | 0.079748 secs |
| 221 | Number of persons by visit start year | 0.076073 secs |
| 225 | Number of visit\_occurrence records by visit\_source\_concept\_id | 0.031793 secs |
| 300 | Number of providers | 0.005537 secs |
| 301 | Number of providers by specialty concept\_id | 0.006073 secs |
| 302 | Number of providers with invalid care site id | 0.006601 secs |
| 325 | Number of provider records by specialty\_source\_concept\_id | 0.006562 secs |
| 400 | Number of persons with at least one condition occurrence, by condition\_concept\_id | 0.018085 secs |
| 401 | Number of condition occurrence records, by condition\_concept\_id | 0.012400 secs |
| 402 | Number of persons by condition occurrence start month, by condition\_concept\_id | 0.084768 secs |
| 404 | Number of persons with at least one condition occurrence, by condition\_concept\_id by calendar year by gender by age decile | 0.067968 secs |
| 405 | Number of condition occurrence records, by condition\_concept\_id by condition\_type\_concept\_id | 0.015999 secs |
| 409 | Number of condition occurrence records with invalid person\_id | 0.014283 secs |
| 410 | Number of condition occurrence records outside valid observation period | 0.019962 secs |
| 411 | Number of condition occurrence records with end date < start date | 0.016532 secs |
| 412 | Number of condition occurrence records with invalid provider\_id | 0.012828 secs |
| 413 | Number of condition occurrence records with invalid visit\_id | 0.048290 secs |
| 420 | Number of condition occurrence records by condition occurrence start month | 0.042020 secs |
| 425 | Number of condition\_occurrence records by condition\_source\_concept\_id | 0.015841 secs |
| 424 | Number of co-occurring condition\_occurrence condition\_concept\_id pairs | 0.026399 secs |
| 500 | Number of persons with death, by cause\_concept\_id | 0.010576 secs |
| 501 | Number of records of death, by cause\_concept\_id | 0.023907 secs |
| 502 | Number of persons by death month | 0.012104 secs |
| 504 | Number of persons with a death, by calendar year by gender by age decile | 0.010606 secs |
| 505 | Number of death records, by death\_type\_concept\_id | 0.007315 secs |
| 509 | Number of death records with invalid person\_id | 0.011667 secs |
| 510 | Number of death records outside valid observation period | 0.009491 secs |
| 525 | Number of death records by cause\_source\_concept\_id | 0.016827 secs |
| 600 | Number of persons with at least one procedure occurrence, by procedure\_concept\_id | 0.048530 secs |
| 601 | Number of procedure occurrence records, by procedure\_concept\_id | 0.034932 secs |
| 602 | Number of persons by procedure occurrence start month, by procedure\_concept\_id | 0.116717 secs |
| 604 | Number of persons with at least one procedure occurrence, by procedure\_concept\_id by calendar year by gender by age decile | 0.144871 secs |
| 605 | Number of procedure occurrence records, by procedure\_concept\_id by procedure\_type\_concept\_id | 0.084539 secs |
| 609 | Number of procedure occurrence records with invalid person\_id | 0.049175 secs |
| 610 | Number of procedure occurrence records outside valid observation period | 0.041149 secs |
| 612 | Number of procedure occurrence records with invalid provider\_id | 0.030878 secs |
| 613 | Number of procedure occurrence records with invalid visit\_id | 0.123940 secs |
| 620 | Number of procedure occurrence records by procedure occurrence start month | 0.110626 secs |
| 625 | Number of procedure\_occurrence records by procedure\_source\_concept\_id | 0.056035 secs |
| 624 | Number of co-occurring procedure\_occurrence procedure\_concept\_id pairs | 0.030142 secs |
| 691 | Percentage of total persons that have at least x procedures | 0.042588 secs |
| 700 | Number of persons with at least one drug exposure, by drug\_concept\_id | 0.076122 secs |
| 701 | Number of drug exposure records, by drug\_concept\_id | 0.045864 secs |
| 702 | Number of persons by drug exposure start month, by drug\_concept\_id | 0.177722 secs |
| 704 | Number of persons with at least one drug exposure, by drug\_concept\_id by calendar year by gender by age decile | 0.256459 secs |
| 705 | Number of drug exposure records, by drug\_concept\_id by drug\_type\_concept\_id | 0.056893 secs |
| 709 | Number of drug exposure records with invalid person\_id | 0.034490 secs |
| 710 | Number of drug exposure records outside valid observation period | 0.041869 secs |
| 711 | Number of drug exposure records with end date < start date | 0.024326 secs |
| 712 | Number of drug exposure records with invalid provider\_id | 0.049273 secs |
| 713 | Number of drug exposure records with invalid visit\_id | 0.095887 secs |
| 720 | Number of drug exposure records by drug exposure start month | 0.109141 secs |
| 725 | Number of drug\_exposure records by drug\_source\_concept\_id | 0.059762 secs |
| 724 | Number of co-occurring drug\_exposure drug\_concept\_id pairs | 0.027198 secs |
| 791 | Percentage of total persons that have at least x drug exposures | 0.057219 secs |
| 800 | Number of persons with at least one observation occurrence, by observation\_concept\_id | 0.021856 secs |
| 801 | Number of observation occurrence records, by observation\_concept\_id | 0.011835 secs |
| 802 | Number of persons by observation occurrence start month, by observation\_concept\_id | 0.034733 secs |
| 804 | Number of persons with at least one observation occurrence, by observation\_concept\_id by calendar year by gender by age decile | 0.061501 secs |
| 805 | Number of observation occurrence records, by observation\_concept\_id by observation\_type\_concept\_id | 0.020443 secs |
| 807 | Number of observation occurrence records, by observation\_concept\_id and unit\_concept\_id | 0.023136 secs |
| 809 | Number of observation records with invalid person\_id | 0.022047 secs |
| 810 | Number of observation records outside valid observation period | 0.016336 secs |
| 812 | Number of observation records with invalid provider\_id | 0.011346 secs |
| 813 | Number of observation records with invalid visit\_id | 0.050968 secs |
| 814 | Number of observation records with no value (numeric, string, or concept) | 0.018383 secs |
| 820 | Number of observation records by observation start month | 0.045173 secs |
| 822 | Number of observation records, by observation\_concept\_id and value\_as\_concept\_id | 0.020773 secs |
| 823 | Number of observation records, by observation\_concept\_id and qualifier\_concept\_id | 0.034998 secs |
| 826 | Number of observation records by value\_as\_concept\_id | 0.020486 secs |
| 827 | Number of observation records by unit\_concept\_id | 0.022459 secs |
| 825 | Number of observation records by observation\_source\_concept\_id | 0.016715 secs |
| 824 | Number of co-occurring observation observation\_concept\_id pairs | 0.019768 secs |
| 891 | Percentage of total persons that have at least x observations | 0.026238 secs |
| 900 | Number of persons with at least one drug era, by drug\_concept\_id | 0.006124 secs |
| 901 | Number of drug era records, by drug\_concept\_id | 0.008420 secs |
| 902 | Number of persons by drug era start month, by drug\_concept\_id | 0.013499 secs |
| 904 | Number of persons with at least one drug era, by drug\_concept\_id by calendar year by gender by age decile | 0.010064 secs |
| 908 | Number of drug eras without valid person | 0.007475 secs |
| 909 | Number of drug eras outside valid observation period | 0.007930 secs |
| 910 | Number of drug eras with end date < start date | 0.011590 secs |
| 920 | Number of drug era records by drug era start month | 0.014005 secs |
| 1000 | Number of persons with at least one condition era, by condition\_concept\_id | 0.020593 secs |
| 1001 | Number of condition era records, by condition\_concept\_id | 0.012011 secs |
| 1002 | Number of persons by condition era start month, by condition\_concept\_id | 0.040717 secs |
| 1004 | Number of persons with at least one condition era, by condition\_concept\_id by calendar year by gender by age decile | 0.047713 secs |
| 1008 | Number of condition eras without valid person | 0.019235 secs |
| 1009 | Number of condition eras outside valid observation period | 0.016707 secs |
| 1010 | Number of condition eras with end date < start date | 0.008313 secs |
| 1020 | Number of condition era records by condition era start month | 0.039554 secs |
| 1100 | Number of persons by location 3-digit zip | 0.020294 secs |
| 1101 | Number of persons by location state | 0.011067 secs |
| 1102 | Number of care sites by location 3-digit zip | 0.008778 secs |
| 1103 | Number of care sites by location state | 0.007360 secs |
| 1200 | Number of persons by place of service | 0.010596 secs |
| 1201 | Number of visits by place of service | 0.036781 secs |
| 1202 | Number of care sites by place of service | 0.010407 secs |
| 1203 | Number of visits by place of service discharge type | 0.017410 secs |
| 1408 | Number of persons by length of payer plan period, in 30d increments | 0.008899 secs |
| 1409 | Number of persons with continuous payer plan in each year | 0.010886 secs |
| 1410 | Number of persons with continuous payer plan in each month | 0.022032 secs |
| 1411 | Number of persons by payer plan period start month | 0.009484 secs |
| 1412 | Number of persons by payer plan period end month | 0.012914 secs |
| 1413 | Number of persons by number of payer plan periods | 0.010689 secs |
| 1414 | Number of persons with payer plan period before year-of-birth | 0.011042 secs |
| 1415 | Number of persons with payer plan period end < payer plan period start | 0.008631 secs |
| 1425 | Number of payer\_plan\_period records by payer\_source\_concept\_id | 0.009480 secs |
| 1800 | Number of persons with at least one measurement occurrence, by measurement\_concept\_id | 0.551826 secs |
| 1801 | Number of measurement occurrence records, by measurement\_concept\_id | 0.212643 secs |
| 1802 | Number of persons by measurement occurrence start month, by measurement\_concept\_id | 0.806123 secs |
| 1804 | Number of persons with at least one mesurement occurrence, by measurement\_concept\_id by calendar year by gender by age decile | 1.044935 secs |
| 1805 | Number of measurement occurrence records, by measurement\_concept\_id by measurement\_type\_concept\_id | 0.213784 secs |
| 1807 | Number of measurement occurrence records, by measurement\_concept\_id and unit\_concept\_id | 0.266165 secs |
| 1809 | Number of measurement records with invalid person\_id | 0.148721 secs |
| 1810 | Number of measurement records outside valid observation period | 0.191973 secs |
| 1812 | Number of measurement records with invalid provider\_id | 0.107827 secs |
| 1813 | Number of measurement records with invalid visit\_id | 0.310181 secs |
| 1814 | Number of measurement records with no value (numeric, string, or concept) | 0.089916 secs |
| 1818 | Number of measurement records below/within/above normal range, by measurement\_concept\_id and unit\_concept\_id | 0.135823 secs |
| 1820 | Number of measurement records by measurement start month | 0.373066 secs |
| 1821 | Number of measurement records with no numeric value | 0.107618 secs |
| 1822 | Number of measurement records, by measurement\_concept\_id and value\_as\_concept\_id | 0.219366 secs |
| 1823 | Number of measurement records, by measurement\_concept\_id and operator\_concept\_id | 0.197007 secs |
| 1826 | Number of measurement records by value\_as\_concept\_id | 0.224814 secs |
| 1827 | Number of measurement records by unit\_concept\_id | 0.279325 secs |
| 1825 | Number of measurement records by measurement\_source\_concept\_id | 0.252593 secs |
| 1824 | Number of co-occurring measurement measurement\_concept\_id pairs | 0.092138 secs |
| 1891 | Percentage of total persons that have at least x measurements | 0.217763 secs |
| 1900 | Source values mapped to concept\_id 0 by table, by source\_value | 0.452636 secs |
| 2000 | Number of patients with at least 1 Dx and 1 Rx | 0.057794 secs |
| 2001 | Number of patients with at least 1 Dx and 1 Proc | 0.029416 secs |
| 2002 | Number of patients with at least 1 Meas, 1 Dx and 1 Rx | 0.193960 secs |
| 2003 | Number of patients with at least 1 Visit | 0.021072 secs |
| 2100 | Number of persons with at least one device exposure, by device\_concept\_id | 0.009018 secs |
| 2101 | Number of device exposure records, by device\_concept\_id | 0.007544 secs |
| 2102 | Number of persons by device records start month, by device\_concept\_id | 0.010899 secs |
| 2104 | Number of persons with at least one device exposure, by device\_concept\_id by calendar year by gender by age decile | 0.015067 secs |
| 2105 | Number of device exposure records, by device\_concept\_id by device\_type\_concept\_id | 0.011868 secs |
| 2125 | Number of device\_exposure records by device\_source\_concept\_id | 0.018650 secs |
| 2200 | Number of persons with at least one note by note\_type\_concept\_id | 0.010255 secs |
| 2201 | Number of note records, by note\_type\_concept\_id | 0.005976 secs |

Query executed in 0.11 seconds