**DOCUMENTATION**

**1.Fact tables**

**1.1.f\_attendence**

Data source: [NHS England Official Statistics: A&E Attendances and Emergency Admissions (www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity)](https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/)

ETL Process: Extraction of documents from the data source in Microsoft Excel 97-2003 worksheet (XLS) and comma-separated-values (CSV) file types. Transformation used the power of the lib Pandas from programming language Python. Firstly, we had to change the file formats of all XLS files due to difficulty to import these sorts of data into PowerBI. We did not transform the data on the years from 2019 to 2023. However, due to failures of data assets in following patterns in the years, we could not consider some variables in the fiscal periods older than 2019-2018 to match the current columns accordingly. These years before 2018 did not have the variable “period”; therefore, we added by extracting the month and year that is included in the filename.

Description:

|  |  |  |
| --- | --- | --- |
| Index | Variable | Type |
| 1 | Period | Categorical |
| 2 | Org Code | Categorical |
| 3 | Parent Org | Categorical |
| 4 | Org name | Categorical |
| 5 | Number of A&E attendances Type 1 | Numerical |
| 6 | Number of A&E attendances Type 2 | Numerical |
| 7 | Number of A&E attendances Other A&E Department | Numerical |
| 8 | Number of attendances over 4hrs Type 1 | Numerical |
| 9 | Number of attendances over 4hrs Type 2 | Numerical |
| 10 | Number of attendances over 4hrs Other A&E Department | Numerical |
| 11 | Patients who have waited 4-12 hs from DTA to admission,Patients who have waited 12+ hrs from DTA to admission | Numerical |
| 12 | Emergency admissions via A&E - Type 1 | Numerical |
| 13 | Emergency admissions via A&E - Type 2 | Numerical |
| 14 | Emergency admissions via A&E - Other A&E department | Numerical |
| 15 | Other emergency admissions | Numerical |

**1.2.f\_population**

Data source:

ETL Process:

Description:

**1.3.f\_diagnosis**

Data source: NHS Official statistics – Hospital Accident & Emergency Activity (https://digital.nhs.uk/data-and-information/publications/statistical/hospital-accident--emergency-activity/2022-23)

ETL Process:

Extraction of documents from the data source in Microsoft Excel 97-2003 worksheet (XLS) and comma-separated-values (CSV) file types. Transformation used the power of the lib Pandas from programming language Python. Firstly, we had to change the file formats of all XLS files due to the difficulty to import these sorts of file into PowerBI. We only extracted the denominated “chief complaints” because lacked data in some historical series. No dataset had the period as a variable, so we had to add according to the year described in the file. In the last period (2023), we extracted and cleaned data from a CSV file to shape into our targeted variable. Date is set as though it has been done on first day of period.

Description:

|  |  |  |
| --- | --- | --- |
| Index | Variable | Type |
| 1 | Code | Categorical |
| 2 | Provider Description | Categorical |
| 3 | SNOMED-CT | Categorical |
| 4 | SNOMED-CT Description\_L1 | Categorical |
| 5 | SNOMED-CT Description\_L2 | Categorical |
| 6 | Total | Numerical |
| 7 | Year | Categorical |
| 8 | Month | Categorical |
| 9 | Date | Categorical |

**2.Dimension tables**

**2.1.d\_date**

Data source:

Not applied.

ETL Process:

Calculated table created inside Power Query on the Power BI software.

Description:

|  |  |
| --- | --- |
| Index | Variable |
| 1 | Date |
| 2 | Year |
| 3 | Quarter |
| 4 | MonthNo |
| 5 | Day |
| 6 | DayMonth |
| 7 | YearMonthNo |
| 8 | YearMonthDay |
| 9 | Month |
| 10 | Month/Year |
| 11 | Quarter/Year |
| 12 | DayOfTheWeek |
| 13 | Day Name |
| 14 | WeekOfYear |
| 15 | DayOfTheWeek |
| 16 | Week/Year |
| 17 | YearWeek |
| 18 | YearMonth |
| 19 | YearQuarter |
| 20 | FiscalYear |
| 21 | FiscalQuarter |
| 22 | FiscalMonth |
| 23 | Future |
| 24 | BusinessDay |
| 25 | Offset Day |
| 26 | Offset Month |
| 27 | Offset Quarter |
| 28 | Offset Ano |
| 29 | Offset Fiscal Year |
| 30 | Complete Week |
| 31 | Complete Month |
| 32 | Complete Quarter |
| 33 | Month/Year Closing |
| 34 | YearMonth Closin |

**2.2.d\_hospitals**

Data source: [NHS England Official Statistics: A&E Attendances and Emergency Admissions (www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity)](https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/)

ETL Process:

Extraction of unique values from the f\_attendence variable “Org Name”. The variables “Lat”, “Long” and “Borough” were defined by researching the physical location of every “Org Name”.

Description:

|  |  |  |
| --- | --- | --- |
| Index | Variable | Type |
| 1 | Code | Categorical |
| 2 | Organization | Categorical |
| 3 | Region | Categorical |
| 4 | Commissioning Region Name | Categorical |
| 5 | Area Team | Categorical |
| 6 | Area | Numerical |
| 7 | Borough | Categorical |
| 8 | Lat | Numerical |
| 9 | Long | Numerical |
| 10 | Code\_1 | Categorical |

**2.3.d\_area**

Data source: London data store

ETL Process: Data curated

Description:

**2.4.d\_diagnosis**

Data source:

NHS Official statistics – Hospital Accident & Emergency Activity (https://digital.nhs.uk/data-and-information/publications/statistical/hospital-accident--emergency-activity/2022-23)

ETL Process:

Extraction of unique values from the f\_diagnosis variable “SNOMED-CT Description\_L1”. The SubCategory variable was developed by the team to cluster data and facilitate the analysis and visualization of the dashboard. The Subcategory created were:

* DERMATOLOGY
* INJURY
* PAIN
* PSYCHOLOGICAL DISTURBANCE
* RESPIRATORY

Description:

|  |  |  |
| --- | --- | --- |
| Index | Variable | Type |
| 1 | SNOMED-CT Description\_L1 | Categorical |
| 2 | SubCategory | Categorical |