Analysis of Nurse Task Duration and Frequency Recorded by RTMS Data

Michael Stroud and Sam Forrest (Advised by Dr. Seo and Dr. Kim)
University of Missouri

Project Motivation

The Electronic Medical Record (EMR) system allows medical personnel to store and access patient data on a computerized system. Unfortunately, using the EMR is very time consuming for medical workers. To find problematic areas of the EMR system, we can analyze tasks that have high variation in average frequency and duration among nurses.

Method

Step 1 To begin analyzing Nurses' EMR data, we first received RTMS data. This data was pulled from the Nurses' EMR equipment as several CSV files.

Step 2 Read this large data set into R by reading each CSV file and binding them to the subsequent file. This creates one file that all data can be pulled from. Filter this data into applicable columns

Step 3 Maintain the entirety data and create a new a dataframe for every nurse that includes only the portion of data collected from that nurse.

Step 4 Sort the dataframes into groups by task. This ensured we can calculate statistics based upon information from each task regardless of when the task occurred.

Step 5 Create dataframes that contains simple statistics describing each group of task. These statistics include frequency, average duration, and variance.

Step 6 Compare entirety data statistics against individual nurse statistics.

Research objective

To identify issues within the EMR system, we must identify potentially problematic tasks. To find these tasks, we will target the tasks that occurred most frequently with high average duration. If a task has high variance in frequency and average duration among nurses, then it may be an indication of an inefficient task. By analyzing these tasks, we can recommend task improvement.

RTMS Data

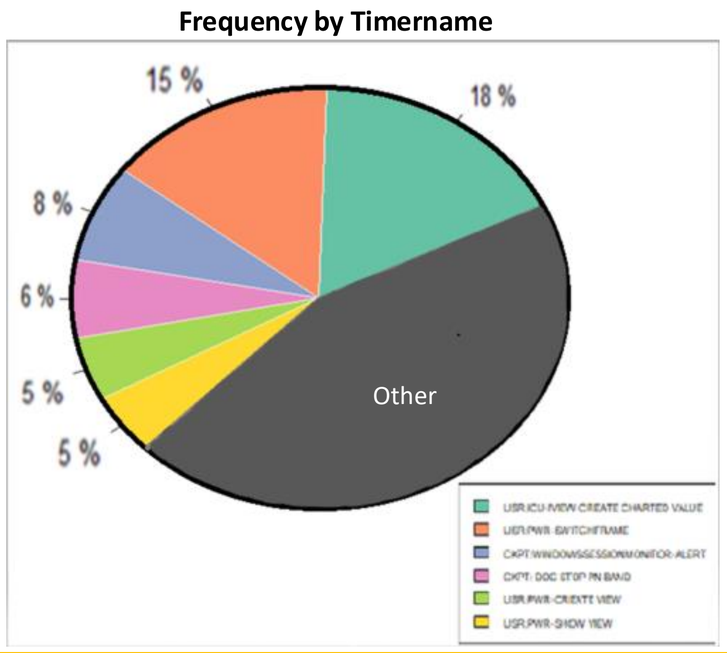
The Real Time Measurement System (RTMS) dataset we received for analysis consists of 160 csv files recorded over a span of just over two months and by 6 nurses. Each csv file is a square matrix consisting of 76 columns and approximately 1 million rows. In aggregate, the RTMS dataset is massive, and a primary goal of our research is to identify ways to get around the practical limitations of dealing with a large dataset.

The primary way that we are reducing the size of the dataset is filtering out unnecessary variables. This greatly reduces the size of the data. Currently we are working on a derivative dataset that contains only 2 of the original 76 columns. The two columns we chose to analyze are timername (task) and duration.

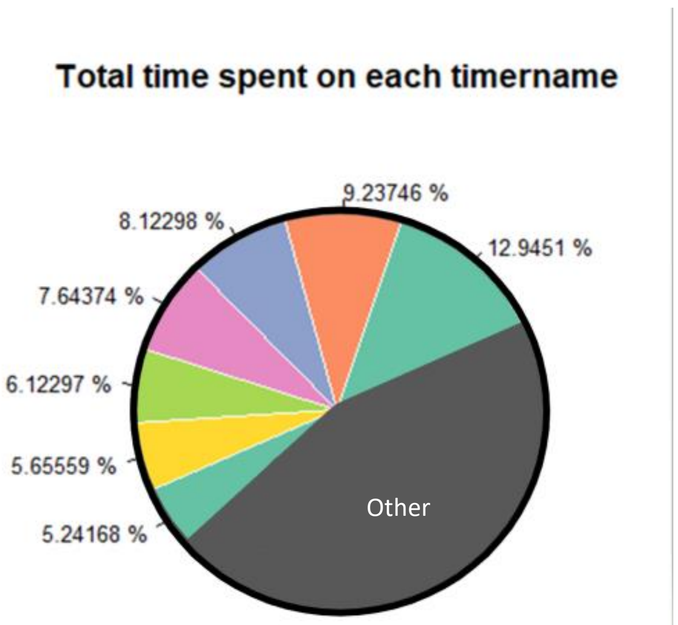
Frequency Analysis

Throughout the RTMS dataset approximately 150 distinct tasks are repeated many thousands of times. We obtained the frequency through grouping. We grouped by task across csv files and found the length of each grouping.

Within the 150 distinct tasks frequency analysis has identified 6 tasks with a frequency exceeding 5 percent.
In total the 6 tasks account for 57 percent of all tasks recorded by the RTMS dataset.



Mean Duration Analysis



The total time spent on each task analysis has identified 7 tasks that account for greater than 5 percent of the total time recorded in the RTMS dataset.

By multiplying the frequency of each task by the average duration of each task we uncover a new metric that can be used to explain the total time spent on a specific task.

- USR:PWR-CREATE VIEW
- USR:ICU-IVIEW-GLOBAL SIGNING RESULTS
- USR: ICU-OPENING-BAND
- ICU: DIALOG TWO STATE GENERIC TIMER
- USR: START CHART
- USR:PWR-OPEN CHART
- USR:PWR-REFRESH VIEW

Duration Variation Analysis