

Analytics at Florida Blue Interview Project Instructions

Introduction

Hello, and welcome to the interview for Sr. Health Business Analyst at Florida Blue!

You should have three separate datasets in .csv format. Below you will find questions relating to each (or a combination) of these files. **Please note that all the health data is entirely fictional and made up by the Analytics team.**

Please complete all items in the **Questions to Complete** section and provide a brief description of your results. You may use any tool to complete the tasks that you wish (R, Python, SAS, Excel, etc.)

Some quick tips:

- Don't be afraid to "dress up" your deliverable
- Context can be useful – Google can be your friend
- Data visualization is encouraged
- If you choose to complete the questions using a programming language, please share your code!

Data Sources

The following consists of the three data sources and their respective definitions. The **ER.csv** file is stand-alone and is not joinable to the other two files. The **healthDat.csv** file is joinable to the **fl_zips.csv** file by the “Zip” field.

ER.csv

Note: This dataset should not be joined to any other dataset, it is stand-alone data

- **Cost:** Total cost by patient in the past year
- **ER_Visits:** Number of visits to the ER by patient in the past year
- **Severity_Level:** The most common severity level associated with the past year’s ER visits

healthDat.csv

Note: All “CC_” fields are flags indicating whether the patient has the condition (1 = yes, 0 = no); Once chronic, always chronic

- **PatientID:** Unique patient identifier
- **Age:** Patient’s age
- **CC_Count:** Number of chronic conditions the patient has
- **HP_Paid:** Total cost by patient in the past year
- **Gender:** Patient’s gender
- **ER_Count:** Count of ER visits in the past year
- **CC_Arthritis:** Arthritis (Osteoarthritis and Rheumatoid)
- **CC_Asthma:** Asthma
- **CC_Atrial_Fibrillation:** Atrial Fibrillation
- **CC_Autism:** Autism
- **CC_Cancer:** Cancer (any type)
- **CC_COPD:** Chronic Obstructive Pulmonary Disease
- **CC_Dementia:** Alzheimer’s Disease and Related Dementia
- **CC_Depression:** Depression
- **CC_Diabetes:** Diabetes (Type I or Type II)
- **CC_Heart_Failure:** Heart Failure
- **CC_Hepatitis:** Hepatitis (Chronic Viral B & C)
- **CC_HIV_AIDS:** HIV or AIDS
- **CC_Hyperlipidemia:** Hyperlipidemia (High Cholesterol)
- **CC_Hypertension:** Hypertension (High Blood Pressure)
- **CC_Ischemic_Heart_Disease:** Ischemic Heart Disease
- **CC_Kidney_Disease:** Chronic Kidney Disease
- **CC_Osteoporosis:** Osteoporosis
- **CC_Schizophrenia:** Schizophrenia and Other Psychotic Disorders
- **CC_Stroke:** Stroke
- **ER_Copay:** Patient’s copay amount for ER visits
- **PCP_Copay:** Patient’s copay amount for inpatient admission
- **IP_Visits:** Patient’s count of inpatient admissions in the past year
- **Zip:** Patient’s zip-code for primary residence

fl_zips.csv

- **City:** City the zip-code resides in
- **County:** County the zip-code resides in
- **Type:** The type of zip-code (standard or PO Box)
- **Zip:** 5-digit zip-code
- **latitude:** Latitude for the centroid of the zip-code
- **longitude:** Longitude for the centroid of the zip-code

Questions to Complete

Question 1: Using the ER.csv file, please describe the relationship between cost and the number of ER visits.

Question 2: Using the healthDat.csv and fl_zips.csv files, please bring in the city and county that each patient lives in and show how you did it. Additionally, if you do not have access to a SQL-supported tool, please write SQL pseudo-code that will accomplish this as well.

Question 3: Using the healthDat.csv and fl_zips.csv files, please provide 2-3 observations or recommendations as they relate to costs. Keep in mind that the goal is to reduce overall costs for this population.

Question 4: Using the healthDat.csv file, please list the 3 most prevalent chronic conditions as well as describe any relationships between conditions that you may have found.

Question 5: Please produce your best dashboard using the healthDat.csv and/or the fl_zips.csv files.
Hint: cost is important!