# Spin #1 Retrospective Document

**Project Scope:**

The project will use baseline BMD, descriptive data, and clinical measurements from the MROS data set to perform a combination of descriptive and predictive analyses, with the aim of understanding and predicting bone fracture incidents.

**Research and Domain questions:**

The project aims to determine the major factors, as gleaned from the baseline BMD and clinical measurements, that are predictive of bone fractures. It will also seek to understand the distribution and patterns of these variables among individuals who have experienced bone fractures.

**Project Goals and Measurable Objectives:**

The goal is to develop an effective predictive model for bone fractures using baseline BMD, descriptive data, and clinical measurements, aiming to achieve a prediction accuracy of at least 80% on the test data set.

**Vision for Final Data Story:**

The data story is targeted at orthopedic surgeons, physical therapists, healthcare administrators, and other professionals involved in the care and treatment of orthopedic patients. Patients could also be an intended audience, where information may be passed to them through a physician to maximize their outcomes. The project will also be valuable for researchers and students in the fields of data science, healthcare, and orthopedics.The findings will be presented as a comprehensive report containing various visualizations and an accompanying presentation highlighting the key findings and implications.

**Team Organizational Structure:**

We want to ensure that all members of the team have an opportunity to fulfill all roles. During each SPiN we’ll be assessing the work that required for the upcoming week over Slack, distributing responsibilities, checking in on Wednesday, and then submitting on Saturday.

**Team Roles and Responsibilities:**

The weekly contributions for this week are:

* Set up and schedule zoom meetings
* Review data source options
* Fill out the Retrospective Document
* Convene with external subject matter expert
* Find and research suitable datasets in the absence of MOI EHR data
* Begin analyzing dataset documentation

**Team Member Accountability and Tracking**

Knowing that we have process updates due and scheduled deadlines already known, it’s incredibly important that we keep communication open and have our scheduled meetings regularly on Wednesday. Up to this point, our group has already had to commit 10+ hours each in training to prepare for our work with potentially sensitive data. Because of this, we’ve had an opportunity to communicate with each other on Slack and already set up ad hoc meetings.

**Schedule and frequency of team meetings**

Team meetings will be held weekly on Wednesday from 7:00PM to 8:30PM with ad hoc meetings scheduled as necessary. We’ve decided to have the meeting scheduled for so long because we’ll plan on meeting with our mentor, Dr. Green, during it.

We also have communication that’ll be occurring throughout the week on our Slack channel - casestudy\_su23\_group03.

**Schedule and frequency of mentor meetings**

Our mentor meeting will occur during our weekly meeting from 7:30PM to 8PM.

**Data storage location, structure, and management**

The raw data flat files will be stored in Juypter Hub but will not be pushed. We’ll then be uploading the data to a Postgres database where it will then be accessed for carpentry, analysis, and our data story.

**Data provenance and assessment**

All team members examined the datasets under consideration for appropriate qualities for this project. Some examples of assessment of the datasets were: datasets through MrOS are able to be joined by participant ID and some datasets have missing values, we hypothesized that they could be due to the non-applicability of certain measures to all participants, data collection errors, or potentially data not being collected or available for certain observations.

**Data acquisition and data sets used**

MrOS – The Osteoporotic Fractures in Men Study funded by NIH. Distributed online via San Francisco Coordinating Center, University of California San Francisco and California Pacific Medical Center Research Institute. Dataset names are B1AUG16, V1FEB23, and FAFEB23. Accessible here: https://mrosonline.ucsf.edu/

**Data carpentry and other coding performed**

Datasets came in SAS file and were imported/analyzed in Python.

**Project work success status evaluated and future work/tasks discussed**

Next steps would include Importing into PostgreSQL and performing initial data carpentry

**Jupyter Notebooks fully documented**

NA

**All paths to data files map to DSA Shared folders**

In progress, but not yet completed

**Current SpIn artifacts are located in TeamArtifacts\SpIn\_1\_Artifacts folder (provide link to first notebook in the pipeline)**

<https://europa.dsa.missouri.edu/user/bromertk/edit/su23CaseStudy_Team03/TeamArtifacts/SpIn_1_Artifacts/README.md>

**Jupyter Notebooks execute without exception**

NA

**Link to Mentor Meeting and key takeaways**

[6-7-23\_Mentor\_Meeting\_Team03.mp4](https://mailmissouri-my.sharepoint.com/:v:/r/personal/jwj8c8_umsystem_edu1/Documents/SU23_DSA8080%20Casestudy/Zoom%20Meetings/6-7-23_Mentor_Meeting_Team03.mp4?csf=1&web=1&e=hDj5H0)

OneNote with meeting minutes and action items: [Mentor Meetings](onenote:https://mailmissouri-my.sharepoint.com/personal/jwj8c8_umsystem_edu1/Documents/SU23_DSA8080%20Casestudy/SU23_DSA8080%20Casestudy/Mentor%20Meetings.one#section-id={3D12AE69-253A-46DF-A3B5-D58BA7327F52}&end) ([Web view](https://mailmissouri-my.sharepoint.com/personal/jwj8c8_umsystem_edu1/_layouts/OneNote.aspx?id=%2Fpersonal%2Fjwj8c8_umsystem_edu1%2FDocuments%2FSU23_DSA8080%20Casestudy%2FSU23_DSA8080%20Casestudy&wd=target%28Mentor%20Meetings.one%7C3D12AE69-253A-46DF-A3B5-D58BA7327F52%2F%29))