**Resources for Writing SOW:**

SOW Template and Guidances from Professor Park: <https://claremontmckenna.box.com/s/cuflzayqby02fzwx73to65tn6o4fvi03>

\* Follow the directions outlined in this document.

SOW Past Example (Flexfit):

<https://claremontmckenna.box.com/s/b698hpb3vegko8by70zil9dhn38l4nkw>

\* We should aim for the thoroughness and brevity (4-5 Page) in this example.

SOW Past Example (Avvo):

<https://claremontmckenna.box.com/s/rtg2cqy9fxwsmqr7jhmrcjyilxmpe157>

East Valley Water District DS Capstone Proposal:

<https://claremontmckenna.box.com/s/miheuh8bnswdb69o8jgsx2nrp6zsho0x>

\* Most of the information we need on SOW can be found here.

East Valley Water District Master Plan 2019:

<https://claremontmckenna.box.com/s/m8m8ptsl7l9ikw0x0jhud0ey4wk5s7es>

\* Focus on the Existing Water System.

Link to Overleaf for LaTeX Editing (Shared with Rishi and Min since it limits editors to 2 people):

<https://www.overleaf.com/6469158384jnyfkpdkwxxt>

**Outline and Draft for SOW:**

* **1 Summary**

East Valley Water District (EVWD, District) is a public rate-based organization that provides water and wastewater services to over 100,000 residents and businesses. Building upon the work from the last semester, the Team will develop water demand prediction and linear programming model along with an user interface designed for regular usage by Operations staff, if possible.

* **2 Problem Statement**

The District is currently dependent on institutional knowledge of key knowledgeable staff and past decisions to adjust flow routing and meet water delivery requirements. As those staff leave the District, new Operations staff need models that predict daily water demand and optimize power usage at each zone and an easy-to-use interface to make optimal flow routing decisions that continue to satisfy customer demands at low cost.

**3 Stakeholders**

Stakeholders in this project are listed below:

* Kerrie Bryan: Director of Administrative Services; overseeing the project
* Rocky Welborn: Water Reclamation Manager; provide data and answer questions regarding flow routing tool
* Residential and commercial customers using water services of the District
* **4 Background**

East Valley Water District (“the District”) is a public rate-based organization in Highland, CA, that provides water and wastewater services to over 100,200 residential and commercial customers. Every day, the District supplies roughly 60 gallons of water per person and over 16 million gallons of water throughout its 30.1 square mile area.

Currently, the optimal decisions to water flow routing are made by key staff with expansive knowledge and years of experience operating the distribution system. When they are unavailable, the District resorts to “do-what-we-did-yesterday” type decisions which usually meet the needs of customers and satisfy the District’s core mission in water provision. As experienced staff leave the District and systematic and environmental conditions change, it is uncertain whether relying on institutional knowledge and past decisions would be sustainable in the long term. The District has requested that the students in the Team develop an easy-to-use tool that can inform and guide new Operations staff, in order to make flow routing decisions in the most optimal, low-cost way.

**5. Technical Background & Proposed Solution:** The Team’s two main goals are accurately predicting water demand in the District and ensuring it is efficiently delivered with minimal power usage. For estimating demand, the amount of water that customers use varies due to environmental factors such as temperature, location, and precipitation, as well as the broader economic and demographic trends. The Team will rebuild the predictive model developed during last semester with improvements in forecasting methods and data used.

For optimizing the power usage during water distribution, the Team will use a linear programming model based on the predictive demand model and relevant inputs. The power required to supply water could vary depending on the elevation because of the differences in hydraulic pressure needed and the distances from pump stations or reservoirs, as well as individual characteristics of substations.

The Team will present both models in a user interface that allows for all relevant inputs and displays key information on projected water demand and power usage minimization process for the district as a whole and for individual zones. The program R will be primarily used in the cleaning of the daily data as well as R/Python for both the development of models.

**5.1 Minimum Viable Product (MVP)**

* **Analysis of Existing Work from Previous Semesters**
  + Analyze R files from student groups working with the District in past semesters
  + Understand both previous intentions of the student groups and present goals of the District
* **Creating Water Demand Prediction Model** 
  + Gather and process historic data needed for the development of the model
  + Explore various factors affecting water consumption to determine relevant variables for forecasting.
* **Outlining Power Optimization Model**
  + Determine all relevant constraints and costs impacting the power usage including but not limited to:
    - Capacity of pumps
    - Predicted demand
    - Number of wells
    - Boosters
    - Drop valves
    - Surface water capacity

**5.2 Post-MVP Features**

* **Creating the Power Usage Optimization Model in R**
  + After outlining the power optimization model, determine the best software development tool to use to run an accurate and efficient model for the District. In addition to technical considerations, ensuring that any backend tools integrate with the user interface will be an important consideration.
  + After choosing a backend software tool, develop the power optimization model and integrate it with the user interface.
* **Integrating Both Models Together:**
  + Connect the demand prediction model with the optimization model to increase efficiency and minimize operator input.

**5.3 Stretch Goals**

* **User Interface for Demand Prediction and Power Optimization Model:** 
  + Integrating both models into an easy-to-use user interface for Operations staff
  + All calculations will automatically run on the back-end and the operators only need to type key input variables to run the models
  + Collecting feedback from the operators to see if there is a way to improve the usability

**6 Risks and Solutions**

* **6.1 Risk #1: Determining Relationship between Variables:** The Team needs to understand which variables will be in the models and their relationships. For instance, if there are discrepancies between geographical regions or different pumps, then our models must regress them separately. The Team is not certain that the relationships between all of the variables that influence the water demand at household level and the power usage of meeting water demand in the District are linear. If all relationships are not linear, then the outputs of the model may not be as accurate as expected.
* **Solution:** Focus on variables with linear relationships and develop a linear programming model that optimizes for the most variables possible.
* **6.2 Risk #2: Changes in Cost Parameters:** In the future, pricing of certain resource costs may change, such as electricity. Using kWh as a metric for cost may be slightly less precise but would be beneficial when including electricity pricing as a variable. Future infrastructure updates could also reduce the amount of electricity used at each pump station and, in turn, affect the cost parameters.
* **Solution:** The Team has already spoken with EVWD to address this issue. We will develop a linear programming model that optimizes for kWh usage. If we have time to address our stretch goals, we will have an advanced inputs section where operators can adjust the kWh usage of different infrastructure locations.
* **6.3 Risk #3: Accessibility for Future Workers:** The employees that use these models after the Team has finished this project must be able to use the model without prior knowledge of the technical aspects of the model.
* **Solution:** To abate this, the Team will build an interface that is simple and easy to use and regularly check with the District to get input on usability.
* **6.4 Risk #4: Data Collection and Cleaning Issues:** The Team might not have access to all the data required for building demand prediction and power usage optimization models. For instance, precipitation and temperature data in the District need to be obtained from third-party sources, which could be unavailable or incorrect. The various formats of datasets could impose additional challenges during data merging and processing.

**Solution:** Ensure that there is clear communication between the Team and the District on data transfer and validation of third-party resources. Regular clarification of data issues should be practiced.

**7 Timeline**

**7.1 Overview of Deliverables**

**Summary of Deliverables:**

**1. Phase 1**

**•Review of existing work from previous semesters**

**•Statement of work on goals and scope of the project**

**2. Phase 2**

**•Cleaned and merged datasets for demand prediction and power optimization model**

**•Outline on power usage optimization model using relevant systemic constraints**

**3. Phase 3**

**•Finalized water demand prediction model**

**•Finalized power usage optimization model**

**•Integration of two models with the existing Excel user interface (if possible)**

**4. Phase 4**

**•Final project report and presentation**

**•Finalized deliverables including comprehensive model package and instruction**

**7.2 Phase 1 (January 18 - February 11)**

During this phase, the Team will work on understanding the main problems the District is facing, establish points of contact, and develop an effective road map for building technical solutions in line with the expectations and needs of the District. By talking to the staff, reading the Project Proposal, and analyzing various resources and data from the District, the Team will gain a better sense of the scope and objectives of the project to clearly define MVP and Post-MVP deliverables.

**7.3 Phase 2 (February 12 - March 15)**

In this phase, the Team will begin collecting and processing the data necessary for rebuilding the predictive demand model and developing optimization model. New consumption and system data from the District will also be used. Appropriate assumptions and processes for the creation of the model will be discussed. The Team will research alternative modeling techniques including machine learning models to improve the accuracy of predictions in the demand model developed last semester. In addition to the demand prediction model, the Team will proceed to outlining the power usage optimization process using linear programming. Determining relevant constraints and an optimal backend development tool that can integrate seamlessly with the existing Excel user interface will be important since the Team will begin developing prototypes of the final product. The Team will also create presentation slides to communicate the progress at the midterm presentation.

**7.4 Phase 3 (March 15 - April 19)**

This is the phase where most of the model development and the integration with an user interface will be done. Maintaining close contact with the District to receive periodic feedback, the Team will work on the demand prediction and power optimization models. If both models are finished earlier than expected, the Team will begin creating a user interface that will integrate them into the Excel user interface so that users only need to input key variables to run the model.

**7.5 Phase 4 (April 20 - May 6)**

The Team will be done with the technical development of demand prediction, cost optimization models, and, potentially, their integration into the existing user interface. During the last phase, the Team will spend the majority of time preparing final reports and presentation decks for both the capstone advisors and the Board of Directors at the District. They will summarize the Team's analysis and include the deliverables for the District with detailed instruction on how to effectively use the models.