Note:-Excel is important for sometimes so please don't disregard it right now, distinct naming is important for models and datasets.

1.Install dataiku from https://www.dataiku.com/dss/trynow/free-edition/

2.After installing follow the instructions for mac from here
https://www.dataiku.com/dss/trynow/mac/

3.How to make the model ?

- 2 ways to build the model, please try both of them:-

1. Give features that you think are important right now that will help in predicting revenue which is target.(Don't recommend focusing on multiple targets like people required, funds needed they mostly acts like bi-products in the equations).

2. If don't know which features are important to give and have less amount of features(also less amount of data(need at least 1800-2000 points and change them on a regular basis)), give all the features.

- After doing this, if you followed the:-

    1. Way, select lab, click on model(prediction) choose your way of doing training, either let the machine decide(Quickly deploy option) or you decide(Expert).
    2. Way, select lab, click on model(prediction), select Expert Training, choose
    -Train/Test Set-
    -Metrics - Determines how much amount of loss you are willing to overlook at (abs sum(MAPE), mean of squares(RMSE,MSE), mean of abs sum(MAE)) or write your own code(have a option for that) for any other errors you want.
    -Features Handling-Which features to take as labels & also choose impute value(replace it with mean,median or a constant value) if you have an empty cell and you don't want to remove that row from analysis
    -Feature generation-For example, if you give Body Weight(BW) and Body Height(BH), you can generate a new one explicitly just like BMI (ex BW/(BH*BH)).Or use the relations (A+B,A-B,A*B) for every feature that you took(Not going to be effective if you choose 3 features it will give 6+6+6(3*(3P2),where P is permutation) extra features which can be computed easily and doesn't cause much effect on time complexity of algorithms)
    -Feature reduction-Not know about every one of them but depends mainly on the problem that you solve(PCA, for example computes nearby distance of most affecting features and takes them into account, basketball players for instance make a jump shot or pass the ball acc. to this).So do a little research and pick one according to the metrics and labels that you gave.

-Algorithms-Do a little research about algorithms that are best for predicting future funds and revenues using the features that you gave.
-Hyperparameters-Read about it in detail from Illustrative guide for Non-tech(Book)
-Runtime Environment-Default settings is fine since mac is fast enough.
-Sample weights-Right now default.

3. After doing this, deploy your model and then predict from laboratory tab. After running predictions, view details.

4.Further, after deploying your models and making predictions, can do 2 things in the long run:-

- If you have made a bad model(the one with less data points) with good accuracy, you can input more data points(please don't input random points) and retrain the same model again and again.(can add more 2-4 features without affecting the model significantly)
- If you want to add more features(at least 20-25), retraining the model won't work in that case, can write 10 lines of code in the future to see the correlation matrix from Hands on ML with scikit-learn and tensorflow

4.Last but not least please know the limitations of this, we can't predict stock prices of everyday by using this techniques (for that you have to write a lengthy code)