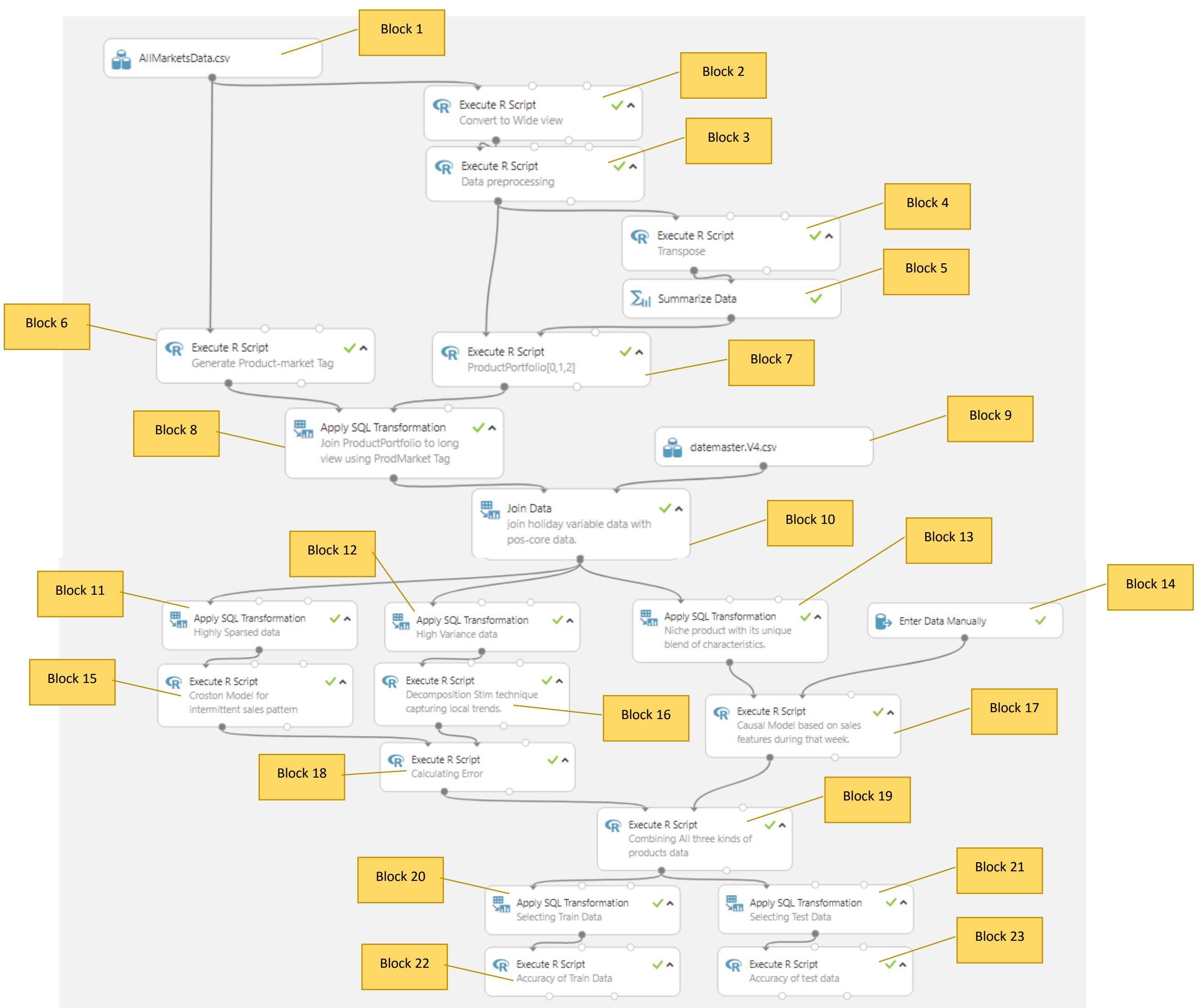
DIGITAL FOUNDRY DEMAND FORECASTING

DF2 – MODEL BLOCK DIAGRAM



DF2 - Blocks Description

Block 1: (Dataset)

- Upload your Data in Azure ML.
- Select the required dataset from "My Datasets" under "Saved Datasets".
- Here, The dataset we are working on has 22 features:

AllMarketsData1.csv

Block 2: (Execute R Script)

- Create Market Product Tag
- Covert required features from long to wide format. In this model MarketProdTags, SALESUNITS, WeekendingDate are converted into wide format with WeekendingDate as Time Variable.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_long_to_wide.R

Block 3: (Execute R Script)

- Data preprocessing: replacing all NA's with 0.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_data_preprocessing.R

Block 4: (Execute R Script)

- Calculate transpose to apply Summarize Data.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_transpose.R

Block 5: (Summarize Data)

- Summarize the data block to generate the summary of the data.

Block 6: (Execute R Script)

- Generate product-market tags.
- Change the Date format

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_product_market_tag.R

Block 7: (Execute R Script)

- Taking preprocessed data from block 3 and Summary features from block 5, calculate sparsity and standard deviation.
- Set product portfolio flag [0, 1, 2] by comparing sparsity and standard deviation with threshold (hard-coded: decided after observing Block 5 result in Azure ML).

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_product_portfolio.R

```
#if std_deviation>100 -------->1 high variance data
#std_deviation<100------------>low variance ----- check for sparsity
#sparsity>25------------------>2 high sparsity
#sparsity<25------------------>0 low sparsity & low variance
```

Block 8: (Apply SQL Transformation)

- Join ProductPortfolio flag generated in Block 7 to long view data from block 6 using ProdMarket Tag.
- t1: Output of Block 6, t2: Output of Block 7

```
select t1.*, t2.ProductPortfolio from t1 join t2 on t1.MarketProdTags=t2.MarketProdTags;
```

Block 9: (Dataset)

- Upload datamaster Data in Azure ML which contains holiday variables flag with respect to weekendingdates.
- Select the datemaster dataset from “My Datasets” under “Saved Datasets”.

datemaster.V4.csv

Block 10: (Join Data)

- Join the outputs of block 8 and 9 using weekending date to obtain original data with product portfolio flags and holiday variables flag.
- Apply Join on WeekendingDate from Block 8 and WeekendDate from Block 9

Block 11: (Apply SQL Transformation)

- Filter highly sparsed data using product portfolio flag.

```
select * from t1 where ProductPortfolio=2;
```

Block 12: (Apply SQL Transformation)

- Filter high variance data using product portfolio flag.

```
select * from t1 where ProductPortfolio=1;
```

Block 13: (Apply SQL Transformation)

- Filtering out data with low sparsity and low variance using product portfolio flag.
- These are Niche product with its unique blend of characteristics.

```
select * from t1 where ProductPortfolio=0;
```

Block 14: (Enter Data Manually)

- Entering the number of weeks to be considered while calculating the discount.
- Here previous 8 weeks data is considered, enter (0, 1, 2, 3, 4, 5, 6, 7, 8).

```
0, 1, 2, 3, 4, 5, 6, 7, 8
```

Block 15: (Execute R Script)

- Execute R block to calculate forecast using Croston Method.
- Separate test and train data by setting “external1” flag 1 for test data and 0 for train data.
- Calculate MASE error for both test and train data.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_croston_model.R

Block 16: (Execute R Script)

- Execute R block to calculate forecast using Decomposition Stlm technique capturing local trends.
- Separate test and train data by setting “external1” flag 1 for test data and 0 for train data.
- Calculate MASE error for both test and train data.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_stlm_decomposition.R

Block 17: (Execute R Script)

- Execute R block to calculate forecast using Causal Model, lm technique.
- Separate test and train data by setting “external1” flag 1 for test data and 0 for train data.
- Calculate MASE error for both test and train data.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_causal_model.R

Block 18: (Execute R Script)

- Bind output from block 15 (Croston model) and 16 (Stlm Decomposition model).
- Add common error column.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_binding_croston_stlm.R

Block 19: (Execute R Script)

- Execute R script block to combine all the three types of forecast outputs
- Add common error column.
- Subset Columns `"PRDCTAG", "MRKTTAG", "error", "external1"`

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_binding_all_datasets.R

Block 20: (Apply SQL Transformation)

- Filter Distinct Train Data from output of Block 19 by checking external1 flag

```
select DISTINCT * from t1 where external1=0;
```

Block 21: (Apply SQL Transformation)

- Filter Distinct Test Data from output of Block 19 by checking external1 flag

```
select DISTINCT * from t1 where external1=1;
```

Block 22: (Execute R Script)

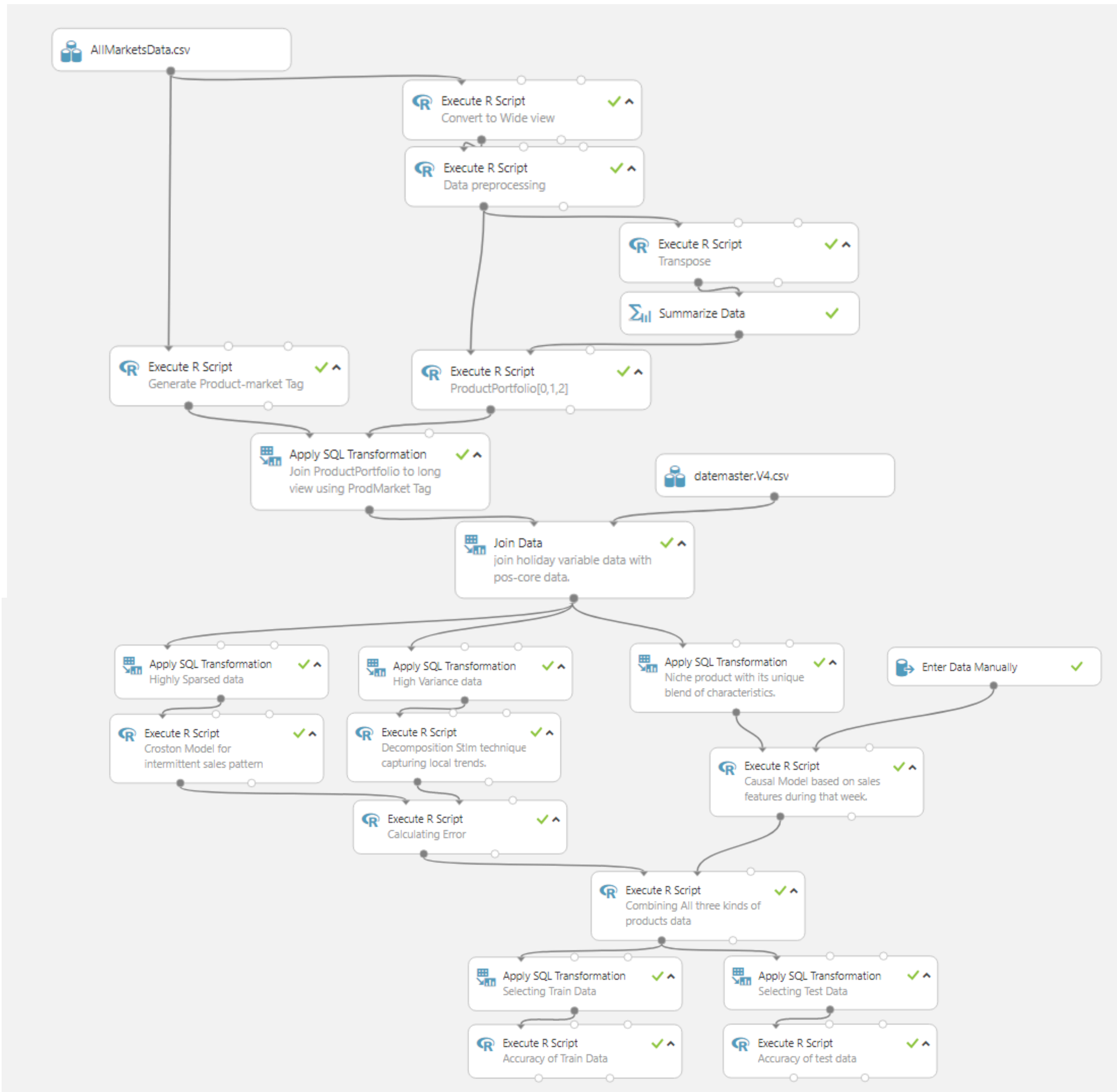
- Evaluate Accuracy of Train data-set from Block 20.
- Reshape to obtain Quality Matrix.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_traindata_evaluation.R

Block 23: (Execute R Script)

- Evaluate Accuracy of Test data-set from Block 21.
- Reshape to obtain Quality Matrix.

https://github.com/kumarchinnakali/digital-foundry-demand-forcasting/blob/master/df2_testdata_evaluation.R


AllMarketsData.csv
Execute R Script
Convert to Wide view
Execute R Script
Data preprocessing
Execute R Script
Transpose
Summarize Data
Execute R Script
Generate Product-market Tag
Execute R Script
ProductPortfolio[0,1,2]
Apply SQL Transformation
Join ProductPortfolio to long view using ProdMarket Tag
datemaster.V4.csv
Join Data
join holiday variable data with pos-core data.
Apply SQL Transformation
Highly Sparsed data
Apply SQL Transformation
High Variance data
Apply SQL Transformation
Niche product with its unique blend of characteristics.
Enter Data Manually
Execute R Script
Croston Model for intermittent sales pattern
Execute R Script
Decomposition Stlm technique capturing local trends.
Execute R Script
Causal Model based on sales features during that week.
Execute R Script
Calculating Error
Execute R Script
Combining All three kinds of products data
Apply SQL Transformation
Selecting Train Data
Apply SQL Transformation
Selecting Test Data
Execute R Script
Accuracy of Train Data
Execute R Script
Accuracy of test data