**Step 2.1 - Model walkthrough (Model setup, Transformation & Model iteration)**

We will be walking you through the modelling process using Papergames data as an example. Since we have already covered the ingestion and classification of data in the training videos, we will be jumping straight to the modelling process.

\*\* We can quickly take them through uploading the file & classifying the data modules\*\* Added “Media\_Demo & Base\_Demo” variables into the data cube.

**Parameter**

In the Modeling module a user is first required to define the model parameters.

1. **Dependent variable -** It is the KPI being modelled. In this example its Revenue sales
2. **Model duration -** The platform in default selects the total time-period ingested in the platform but we also have the option of selecting a more recent time-period (Do select May 1st 2021 to Oct 31st 2021)
3. **Model type -** It has one option
   1. **Unpooled -** When the data is available only at a single dimensional level. We are only presented with this option. Which in our case in this example.
4. **Model form –** It has 2 options
   1. **Additive -** For this example we have selected the additive form
   2. **Multiplicative –** It is recommended when the KPI has high seasonality and/or when the media tactics are expected to have synergy effect between them. It is worth mentioning that the multiplicative form is a bit more complex than the additive form, however the platform takes care of this and provides us results like an additive model

**Correlation**

1. It allows us to check for correlation between the KPI & the independent variables and check for multicollinearity among the independent variables themselves
2. The platform also offers the options of selecting the time-period for which the correlation check is required and a feature to set a threshold for correlation %, where the platform highlights a tactic which crosses the set threshold

**Create transform measures**

1. This tab contains a list of transformations usually used in MMM (Marketing Mix Modeling)
2. It is a simple click and transform kind of an approach, which does not involve any manual calculation
3. We simply pick the transformation of interest and select the tactic to the transformed
4. We can look at this in detail, with an example, as we build a model shortly

**Switch to Auto mode**

1. Switching to this mode enables the platform to create multiple Ad-stock transformations of a/multiple variables and returns the transformation with the highest correlation with the KPI
2. It saves time on manually creating transformations; however, the platform only applies Ad-stock transformations now

**Variable selection**

1. Scrolling down further we have 2 panes, the pane on the left contains the list of variables ingested in the platform and the pane on the right is where the variables to be tested in the model are moved to. **(Click on Updates)**
2. To select the variables, we shall simply check the variables of interest and move to the selected measures area (right pane) or search for the variables in the search bar (on top of the left pane) and move it to the selected measures area
3. The list of variables in the selected measures pane will be tested in the model

**Priors**

1. Before running the model, we have the option of setting priors I.e., informing the model of empirical evidence of relation between the independent variables and the KPI.
2. We can provide co-efficient and standard deviation to a tactic(s). This will give us the out of interest, contribution, which we could use to calculate performance measures like effectiveness and ROI (Return on Investment).
3. We shall look at this in detail, with an example, while we build the model

**Model output Page**

The model output page displays the results of the given iteration. Let us look at features available in it.

**Select model –** It is a list of all the saved and unsaved iterations of the model. We shall scroll through it to view any iteration of interest for future reference. To save an iteration we can simply provide a name in the “**Enter name to save model**” section and click on save.

Below we have the various tabs of results in the model output page. We shall discuss each of them in detail one after the other.

**Model fit**

1. **Model fit chart –** It consists of the predicted sales (red line), modelled sales (orange line) and blue columns representing residual (error i.e., difference between predicted and modelled sales). The chat also displays the model stats (RSQ & MAPE- Mean Absolute Percent Error), where the **acceptable range for a daily model is RSQ > 80% & MAPE < 20%**
2. **Model fit data –** It consists of the results a variable level, where the list of variables selected to be tested in the model, their respective coefficients along with its statical validation measures are displayed
   1. **Standard error –** it is a measure of uncertainty. The higher the value, the more uncertain is the coefficient of a given variable
   2. **T Stat -** A T-stat of +2/-2, depending on the nature of the variable, indicates that the coefficients are significant
   3. **P value –** A P value of <=0.05 indicates that the coefficients are significant
   4. **VIF (Variable Inflation Factor) - it** is a measure to indicate multi-collinearity among the independent variables. A value above 25 highlights the presence of multi-collinearity.

**Model iteration**

1. Now that we have viewed the results of the selected base variables, we shall try a**dding a media variable** and see how the model reacts.
2. **Update Model –** In order to edit or add new variables to a current iteration a user must click on the **update model** button that is available on the top right of the model output window. This takes us back to the model setup window.
3. **Adding media variable –** A raw media variable cannot be included in the model because the carryover impact will be accounted for in the model. To apply that lets transform the media variable **(FB\_Ap\_June)** 
   1. **Transformation –** As mentioned earlier, select the transformation of interest **(select gamma)** and the variable to be transformed and click on continue.

\*\* We can show them the slide we created explaining about the transformation; File Name < Saas Training Model walkthrough> 3rd & 4th Slide. \*\*

Going back to the Platform.

* 1. As discussed, here, we input the parameters for transformation. In the case of gamma transformation. There are 3 parameters,
     1. **Decay –** Indicates the % of impact that is being carried over to the next day **(input 0.2)**
     2. **DOF (Degree of Freedom) - Provide** the build aspect to the transformation, for example, a DOF of 1 indicates that the impact is immediate, whereas when the DOF is 2 it indicates that the real/majority of the impact occurs from the 2nd day **(Input 1)**
     3. **Week -** (days in this case, as it is a daily model) indicates the period for which the impact is allowed to carryover to. For example, if week/day is 5 the carryover impact stops at the 5th day **(Input 7)**
  2. The platform also allows us to rename the transformed variable. It is suggested to add a meaningful suffix to the input name. For example, Gamma\_217. Indicating gamma, decay, DOF & week/day respectively.
  3. There is an additional option of creating multiple transformations at the same time. By clicking on the “+” sign and adding as many new transformations as possible for the selected variable.
  4. To generate a new transformed variable, a user must click on save and only then click on generate measure.

**Input new transformed variable**

1. Now that we have created the new transformed variable. We shall go to the model output page and click on update model to select the new variable to test it in the platform & run the model including it
2. \*\*\*Mention how the RSQ & MAPE have changed and proceed to adding another media variable\*\*\*
3. Select **YT\_Ap\_June 🡪** apply gamma transformation decay (0.3), DOF (1), Weeks (7)
4. **\*\*\***Talk about how the RSQ & MAPE and highlight how the **(FB\_Ap\_June)** media variable has taken up a negative coefficient\*\*\*
5. **\*\*\***Remove YT data & highlight how FB has taken up a positive coefficient now but also highlight that the RSQ has decreased\*\*\*
6. Now we shall use the positive coefficient of FB as its prior and give it a 20%(0.031) standard deviation & also include the YT tactic along with it.
7. Mention how the RSQ has improved from the previous iteration and both FB & YT have taken up a positive coefficient\*\*\*

This is one way of arriving at a coefficient and these coefficients could be validating the contribution % with the spend share % of each of the tactic and make sure that it makes business sense along with a good model fit.