**Data Prep for RedPoint POC**

* Data is stored in hive in mktgplatformdb
* ddl’s, hql’s, sh, and mapping spreadsheet are on the linux server - /home/hdp\_oc\_team/RedPoint
* hadoop infrastructure created the following db’s:
  + - * + fortknox
        + godaddyprivacyapp
        + godaddyinstorecredit
* There appears to be a problem loading data to all of the tables (except fortknox. fortknox\_shopper\_snap)
* FYI – replication is not yet fully restored to fortknox. dbo.fortKnox\_shopper on TRO2, but data is up to date.
* Gdshop\_marketing\_contact – table and hive script created, but needs data from privacy app. Will need to reload.
* Created columns for in-store credit in shopper table, but waiting for data. Unable to create the underlying process to summarize until the data is available. Will complete when I return.
* Created a shell file to load all the tables from the individual hive scripts - /home/hdp\_oc\_team/RedPoint/redpoint\_load.sh
* Tried to maintain consistent naming standards in the tables and columns.
* Tried to maintain consistent data types in the columns.
* Encrypted email is not available
* Do not contact data is out of date – need a new source
* Did not create a table for Visit Page Scores
* Did not create a table for scoring models
* Did not create a table for no paid/free products – might have it done by end of day.