Documentation of Assumptions:

We have made several general assumptions given the time constraint.

**Acquisition Date** – NA’s have been removed, because we do not know the date. We are removing Acquisition dates before September 30th, 2016, because the agent leads table does not have observations further back then that. We remove all records with acquisition date before the first lead date time, because the acquisition can’t happen before the first lead.

**Lead Type and Vendor** – With 76 unique names in the LeadVendor column we had to make general assumptions of where they belonged given the marketing information given to us. We groped them into the lead channels given to us with the best of our ability. There were several LeadVendors we did not know where belonged and ultimately decided to place them where we assumed, they belonged or into a unknown group. We also assumed that if lead type is “Organic” the vendor must be organic. Same for email, where vendor must be email if lead type is email.

**REAgentID** – Some REAgentIDs had what we believed to be too many unique observations and decided to remove the once over 200. Again, a conservative estimate of removal, but we assumed that lower than that could be possible. Over 200 seems very unlikely and must either be errors or test REAgents.

**TalkTimeMinutes** – We have removed every observation where talk time was either over 60 or 0.0. Anything above 60 minutes seems highly unreasonable and could either be errors or other misreporting. From a general perspective it makes sense for Sales Representatives to not take up too much time with each customer. Sales meeting might be a better option if phone conversations are running really long.