1. **Business Questions**:

How do promotions in different display locations (e.g., in-store, mailers) impact customer purchase behavior?

1. **Datasets**: <https://bradleyboehmke.github.io/completejourney/articles/completejourney.html>

* Coupon redemptions
* Transactions
* Products
* Promotions
* *Data Assumptions:*

Other than the FK (foreign keys) and PK (primary keys), we assume that the redemption\_date on coupon\_redemptions dataset is the same as transaction\_timestamp on transactions dataset.

1. **Hypotheses**:
2. Customer is likely to buy stuff at store-front even without coupons.
   * Data need:
     + Outer join transactions dataset and coupon\_redemptions dataset based on household\_id and might be transaction\_timestamp. This will result in transactions without coupon redemptions => called this dataset: transactions\_no\_coupon
     + Join the transactions\_no\_coupon dataset and promotions dataset to get the display\_location and mailer\_location for products, might be joined based on product\_id => called this dataset: trans\_no\_coupon\_products
     + Filter for products having display location is store-front, then calculate the total sales grouping by product\_id from trans\_no\_coupon\_products dataset.
   * *Note*: make sure to have another column in transactions dataset which include date only (ex: 2017-05-11), transaction\_timestamp has both date and time
3. Stuff at certain display\_location with coupons will have the most purchase.
   * Data need:
     + Inner join transactions dataset and coupon\_redemptions dataset based on household\_id and might be transaction\_timestamp
4. Customer is likely to redeem coupons if products are in mail (mailer\_location) at interior page feature.
5. Customer doesn’t use the coupons if products are not shown in mail (mailer\_location).
6. **Questions for TA:**
   1. Discuss question
   2. coupon\_upc: Uniquely identifies each coupon (unique to household and campaign)
      * Some coupons are redeemable for multiple products.