1. **Business Questions**:

How do promotions in different display/mailer 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).

# Data explanation

1. **Final\_Transactions\_With\_Redemptions.csv**

* Including every transaction having coupons redeemed
* No unique key included
* No total\_sales calculated

1. **Transactions\_W\_Redemptions\_Display\_Location.csv**

* Unique key combination: product\_id, store\_id
* Including total\_sales of each unique combination of product\_id and store\_id, along with other info about department, product\_type, product\_category
* Only have display\_location != 0
* Display location (0-no display, 1-store front, 2-store rear, 3-front end cap, 4-mid-aisle end cap, 5-rear end cap, 6-side aisle end cap, 7-in-aisle, 9-secondary location display, A-in-shelf)

1. **Transactions\_W\_Redemptions\_Mailer\_Location.csv**

* Unique key combination: product\_id, store\_id
* Including total\_sales of each unique combination of product\_id and store\_id, along with other info about department, product\_type, product\_category
* Only have mailer\_location != 0
* Mailer location (0-not on ad, A-interior page feature, C-interior page line item, D-front page feature, F-back page feature, H-wrap front feature, J-wrap interior coupon, L-wrap back feature, P-interior page coupon, X-free on interior page, Z-free on front page, back page, or wrap)

1. **Transactions\_W\_Redemptions\_Not\_Displayed.csv**

* Unique key combination: product\_id, store\_id
* Including total\_sales of each unique combination of product\_id and store\_id, along with other info about department, product\_type, product\_category
* Only have display\_location == 0
* Display location (0-no display, 1-store front, 2-store rear, 3-front end cap, 4-mid-aisle end cap, 5-rear end cap, 6-side aisle end cap, 7-in-aisle, 9-secondary location display, A-in-shelf)

1. **Transactions\_W\_Redemptions\_Not\_Mailed.csv**

* Unique key combination: product\_id, store\_id
* Including total\_sales of each unique combination of product\_id and store\_id, along with other info about department, product\_type, product\_category
* Only have mailer\_location == 0
* Mailer location (0-not on ad, A-interior page feature, C-interior page line item, D-front page feature, F-back page feature, H-wrap front feature, J-wrap interior coupon, L-wrap back feature, P-interior page coupon, X-free on interior page, Z-free on front page, back page, or wrap)