🡪 Members Dataset

* Important columns:
* Cohort\_month
* membership id
* Membership\_type\_code or membership\_type\_desc
* Join\_date
* Tenure\_grp
* Renew\_ind
* Last\_renew\_date
* Renew\_date
* Next\_renew\_date
* Plus\_status\_before\_ren
* Plus\_status\_after\_ren
* Auto\_renew
* All the below demographic variables
* Membership\_type\_code and membership\_type\_desc (chose the description column over the code?)
* Interesting Questions in this dataset (how much each of the below factor play a role in customers’ renewal decision-making process, demographic analysis)🡪
* Distance to the club (Miles\_to\_club column)
* Age (hhh\_age\_desc)
* Marital status (marital\_status\_desc)
* Income (income\_desc)
* Household size (hh\_size\_desc)
* Number of children (Nbr\_children\_desc)
* Ethnicity (ethnic\_desc)
* How many of the total members renewed? Base to Plus? Base to Base? Plus to Plus? Plus to Base?
* Is there a particular segment of customers that has higher renewal rates?
* Is there a correlation between the income/miles to club and renewals/sales?
* **Are the members who are on auto-renewal tend to renew their membership more frequently than the contrary**🡪 **“autorenew\_ind”**
* **Which consumer type is more popular: savings or business? (Membership\_type\_code, Membership\_type\_desc)-->The split of savings & business in terms of sales & renewals**

🡪pos dataset

* MEMBERSHIP\_ID
* VISIT\_DATE
* VISIT\_NBR

🡪Join **member** table with **pos** on membership\_id

* Understand how members engage in their membership journey (how consistent they are visiting, how much they buy and what products, etc.)
* How consistent members are visiting?
* Members’ number of visit🡪visit\_nbr

How do the renewal rates differ by clubs/regions?

Category

* What are some of the top categories? (by sales)
* Top categories by Regions?
* What are the top categories (maybe at DMM or GMM level) for renewing customers?

Visit Frequencies

* On an average how many times (a month/ a year) does a FY renewed member visit vs unrenewed?
* Weekend vs Weekday comparison?
* Time of the day comparison?
* How often do they visit? – inter-visit frequency? Tricky to calculate

Sales

* What is the split for sales among base and plus members?
* **What is the difference in sales/profit for renewing/non-renewing customers?**

🡪 Join **GMM\_DMM** with **pos** on CATEGORY\_NBR

* Important columns in Category/ **GMM\_DMM**: CATEGORY\_NBR, GMM, DMM, Category (chose description over the code, good idea?)
* Note: GMM is a broader umbrella of DDM, which is a broader umbrella for category
* Merchandise Hierarchy: GMM -> DMM -> Category -> Sub-Category -> Item (The items and their associated sub-category and category are provided in the POS table)
* Interesting Questions?
* what products members buy – most popular in general and which is the most popular among specific demographic groups?
* Compute profits: Profit = Sales - (Unit\_qty\*Unit\_cost)
* Profits Analysis:
* Break profits by membership type (plus/basis)
* Break profits by demographic groups (will require to join the members\_pos table)

🡪Join **tender** and **pos** on VISIT\_NBR

* Shall we join both tables on CLUB\_NBR as well?
* TENDER\_TYPE & TENDER\_TYPE\_DESC (the first one is a code for the payment method that the members used in their visit and the later specifies it type: visa, cash, shopping card, American express, etc.)
* Calculate the sum of TENDER\_AMT foe each tender type “TENDER\_TYPE\_DESC”

🡪Goals to accomplish today:

* create a data map to join the tables
* come up with the business questions that will be answered thru each join (make sure we’re in alignment with the objectives)
* identify the columns that will be included in the join
* discuss the appropriate data type conversion for each column
* *We can keep digging in the data forever for the analysis, what will add more value to the company is in the way we present and display our analysis – visuals and dashboards are extremely important*