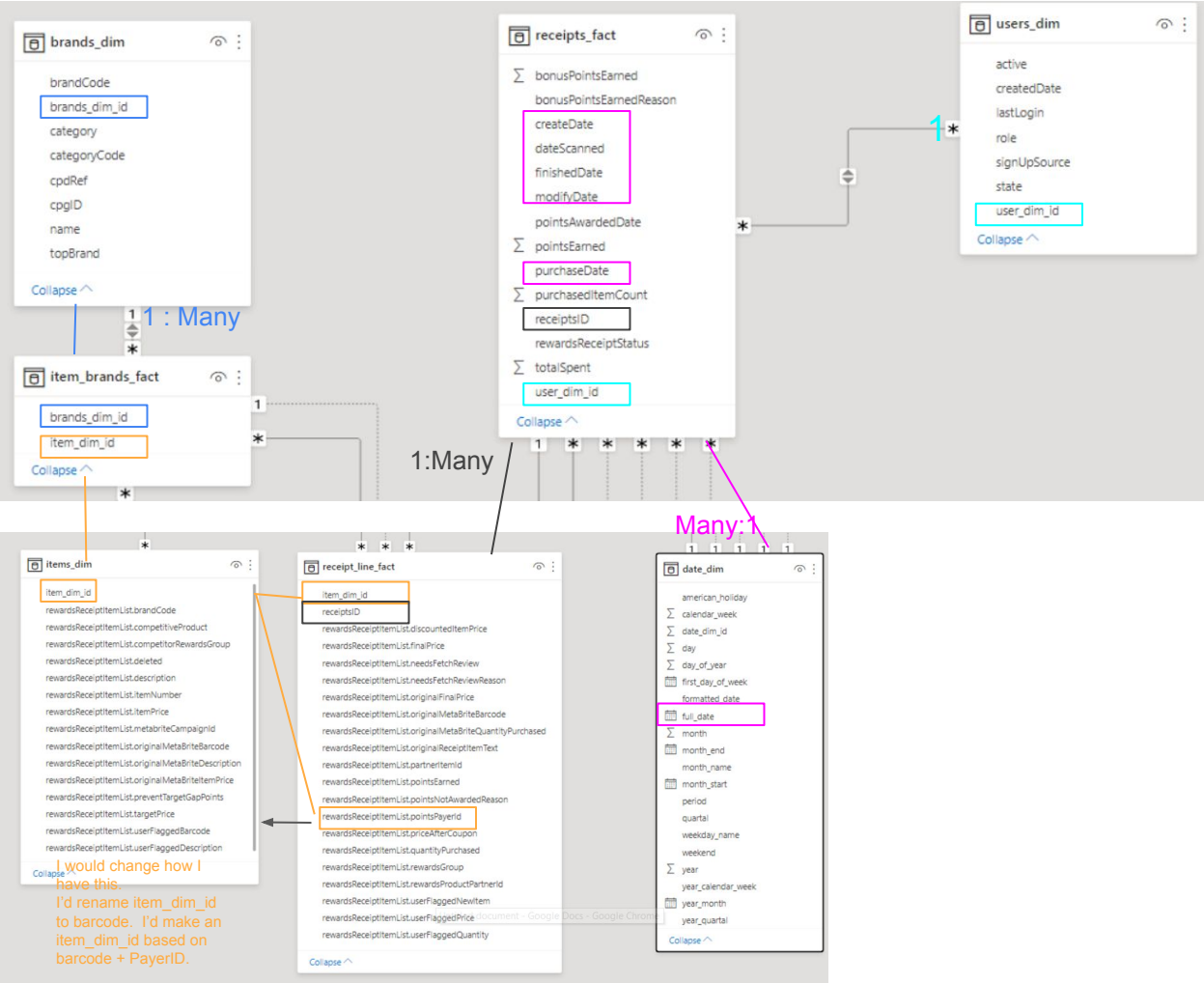

For a data warehouse, I would use a fact and dimension structure.

I would need to know much more about the data to properly model it. I would call this a draft.

I struggled with the Brands file and its associations. All of the barcodes on that file start with 51111 but very few barcodes on the receipts file have that barcode. So, I'm unsure if that's a key. I know that barcodes can have many brands but there are several fields on the receipts file that could link the two tables. I made an assumption that payerID is related to a brand.

I didn't use the brands file at all in the analysis because when I joined it to receipts, I got very few or no results. This makes me think I'm not using the right fields.

The user_dim table actually has some transactional data points (login date) so I would rethink how I did that. Also, there seems to be duplicate data.

I converted dates to UTC.

I used color coding to show joins