**Project Update: Receipt Rewards Data Warehouse**

Hi Team,

I've completed the initial analysis of our rewards data and wanted to share key findings before our meeting.

**Data Model**

I've designed a streamlined relational model with four core tables:

* **BRANDS**: Product information with categories
* **USERS**: Customer profiles and activity data
* **RECEIPTS**: Transaction records with status and totals
* **RECEIPT\_ITEMS**: Individual purchased products

This model supports all our analytical needs while simplifying the current complex JSON structure.

**Business Insights Ready to Deliver**

The new structure will enable us to answer critical questions:

* Top 5 brands by receipt volume (monthly trending)
* Spending patterns for accepted vs. rejected receipts
* Brand preferences among recently acquired users

**Data Quality Concerns**

Several issues need addressing before implementation:

1. **Duplicate users**: 70 user IDs appear multiple times (283 duplicate records)
2. **Data completeness**: Key fields like brandCode missing in 20% of brands
3. **Extreme values**: Total spent ranges from $0 to $4,722 with outliers
4. **Geographic concentration**: 90% of users from Wisconsin

**Questions for Discussion**

1. Which source system should be the user data authority?
2. Should we exclude outlier transactions above a certain threshold?
3. What date range of historical data needs to be migrated?
4. How often should we refresh the warehouse?

Let me know if you'd like to review any specific aspects before our meeting.

Thanks,

P