**What is MailChimp?**

Mailchimp is the world’s leading marketing automation platform for small businesses. To this end, the platform offers services including marketing automation, landing pages, email templates and product recommendations (affectionately known as P-REX).

**MailChimp’s goals are to publish the right content to the right person at the right place at the right time.**

---

**What are Personalized Product Recommendations?**

Using the purchase history of each customer to make smart, data-driven predictions about what they’ll want to buy in the future.

Our 1st few weeks were reviewing customer feedback about the existing system, understanding pain points, and seeing if there were ways we could improve the existing P-REX system.

---

**Central Business Question: Can we improve the relevance of P-REX for consumers who are the recipients of Product Recommendations from MailChimp customers?**

---

**Datasets**

**Raw Data:**
- Sample of ~1,000 stores
- Historical transactions for 3 years
- Product details, including text descriptions

**Cleaning and Processing:**
- Removed NA’s, aggregated sales for the same customer, and same products
- Transformed the datasets into user * product matrices

**Central Business Question:** Can we improve the relevance of P-REX for consumers who are the recipients of Product Recommendations from MailChimp customers?

---

**Recommendations: Business Impact:** Expansion of the P-REX feature will give MailChimp’s customers a greater ability to grow their small business by using personalized e-commerce tailored to their consumers.

---

**Testing and Results:**

We tested cosine similarity, which is MailChimp’s current method, and SoftImpute on small and medium stores*. The metrics we used to train the model and tune λ is NMSE:

\[
\text{Normalized Mean Square Error (NMSE)} = \frac{\sum (\text{Model Predictions} - \text{Observed Ratings})^2}{\text{Observed Ratings}^2}
\]

With the optimal λ, we masked 20% of the purchase matrix and tested recommendations using Hit Rate @ 3: how many items the model can detect as being purchased i.e. the top 3 items likely to be purchased by the user.

We have run the Soft-Impute methodology over 74 small stores and 112 medium stores using a stratified sample.*

For both the small and medium stores, Soft-Impute outperforms the Cosine Similarity recommender system, but the difference is only statistically significant for small stores.

---

**Recommendations: Business Impact:** Expansion of the P-REX feature will give MailChimp’s customers a greater ability to grow their small business by using personalized e-commerce tailored to their consumers.

---

Thanks for a wonderful summer in Atlanta, Georgia!