# DEMAND FORECASTING @ EUROGROCER MSiA 410

Professor Joel Shapiro





A. How valuable would it be to the business if it could completely eliminate **stock-outs** arising from incorrect demand prediction?

(Hint: using the info only on p6-7 what is the stock-out rate, and thus, the total foregone profit, based on the overall revenue and gross margin?)

B. What about **spoilage**?

(Hint: when you have extra stuff, there's no demand for it, so it's about avoiding cost, not foregoing profit...also only using info on p6-7)

**New Question:** how does margin relate to relative impact of stock-outs versus spoilage? That is, if you had a high-margin good, which is worse - a stock-out or spoilage? How about the reverse?

A. How valuable would it be to the business if it could completely eliminate **stock-outs** arising from incorrect demand prediction?

(Hint: using the info only on p6-7 what is the stock-out rate, and thus, the total foregone profit, based on the overall revenue and gross margin?)

(3.4% stock-out rate) x (\$8B revenue) x (34.2% margin) = **\$93M** 

That is, if they had this 3.4% of stock on shelf, then they would have made a 34.2% margin on it.

B. What about **spoilage**?

(Hint: when you have extra stuff, there's no demand for it, so it's about avoiding cost, not foregoing profit...also only using info on p6-7)

(5.8% stock loss) x (50% due to spoilage) x (\$8B revenue) x ((1 - 34.2% margin) = \$153M

For any given item that spoils, the cost to the store was could have been saved.

Cost = retail price \*(1-34.2%)

E.g., if an avocado retails at \$2, cost to the store = \$2 \* (1-profit margin = %) = \$2 \* (1-34.2%) = \$1.32

A. How valuable would it be to the business if it could completely eliminate **stock-outs** arising from incorrect demand prediction?

(Hint: using the info only on p6-7 what is the stock-out rate, and thus, the total foregone profit, based on the overall revenue and gross margin?)

(3.4% stock-out rate) x (\$8B revenue) x (34.2% margin) = **\$93M** 

B. What about **spoilage**?

(Hint: when you have extra stuff, there's no demand for it, so it's about avoiding cost, not foregoing profit...also only using info on p6-7)

(5.8% stock loss) x (50% due to spoilage) x (\$8B revenue) x ((1 - 34.2% margin) = \$153M

**New Question:** how does margin relate to relative impact of stock-outs versus spoilage? That is, if you had a high-margin good, which is worse - a stock-out or spoilage? How about the reverse?

A. How valuable would it be to the business if it could completely eliminate **stock-outs** arising from incorrect demand prediction?

(Hint: using the info only on p6-7 what is the stock-out rate, and thus, the total foregone profit, based on the overall revenue and gross margin?)

(3.4% stock-out rate) x (\$8B revenue) x (34.2% margin) = **\$93M** 

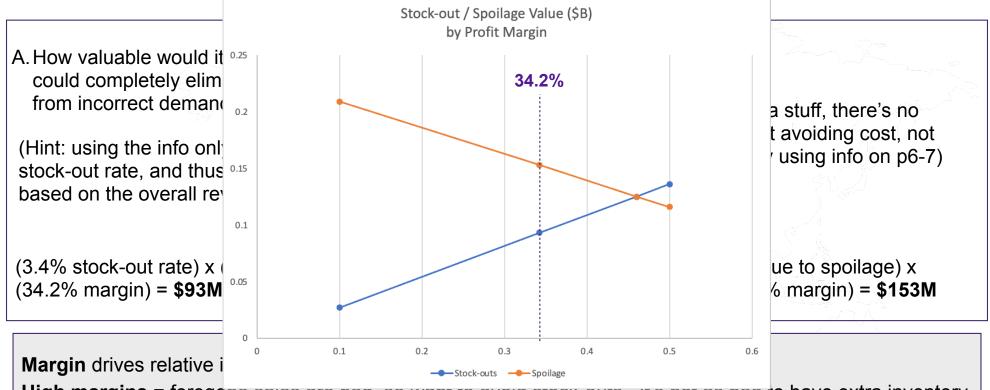
## B. What about **spoilage**?

(Hint: when you have extra stuff, there's no demand for it, so it's about avoiding cost, not foregoing profit...also only using info on p6-7)

(5.8% stock loss) x (50% due to spoilage) x (\$8B revenue) x ((1 - 34.2% margin) = \$153M

Margin drives relative impact of stock-outs or stock-loss on bottom line.

**High margins** = foregone sales are bad, so want to avoid stock-outs. It's not as bad to have extra inventory. **Low margins** = foregone sales aren't so bad - would rather have foregone sales than extra inventory.



High margins = foregone sales are pag, so want to avoig stock-outs. It's not as pag to have extra inventory.

Low margins = foregone sales aren't so bad - would rather have foregone sales than extra inventory.

# Q2. IMPACT OF PREDICTION IMPROVEMENT

## 2. By product category

- A. Cost of stock-outs and spoilage for each product category
- B. Beneficial impact on profit for 1% prediction improvement
- C. Total impact on profit

See Excel spreadsheet for calculations

## 3. A) Upfront labor cost of IBM team and Eurogrocer team

Upfront labor costs (planning, setup, development of ML solution) =

FTE cost / hour \* # of hours

6 people 4 people

\$40k / month \$200k / year = \$16.67k / month

100% utilization 80% utilization

\$240k / month \$53.3k / month \$853 / month

**SUM = \$294.2k / month** 

\$200k / year = \$16.67k / month

9 hours worked / month

22 days / month @ 8 hours / day

**TOTAL OVER 6 MONTHS = \$1.765M** 

#### 3. B) Expected annual labor cost for Eurogrocer team

Note there are ongoing fixed costs and ongoing variable costs

#### Maintenance and operation of models

Fixing bugs and data issues, filtering / cleaning data, data gathering, model upgrades

Other costs, such as subscription costs / hardware leasing treated as negligible here

## Training / running models, data storage

Depends on # of models built and maintained

This is a subjective decision - maybe ideal to have as many models as there are SKU-store combos?

This is probably inefficient: likely some commonalities between similar products or products in a single store.

Here, end decision was 1 model for each product category (e.g., "Bakery") for all stores

#### 3. B) Expected annual labor cost for Eurogrocer team

Note there are ongoing **fixed** costs and ongoing **variable** costs

#### Maintenance and operation of models

Fixing bugs and data issues, filtering / cleaning data, ————data gathering, model upgrades

Other costs, such as subscription costs / hardware leasing treated as negligible here

## **Development effort**

4 people

\$200k / year

30% utilization

= \$240k / year

#### **Subject matter experts**

\$200k / year = \$16.67k / month

22 days / month @ 8 hours / day

2.5 hours worked / month

= \$2841 / year

**SUM = \$242.8k / year** 

#### 3. B) Expected annual labor cost for Eurogrocer team

Note there are ongoing **fixed** costs and ongoing **variable** costs

## Training / running models, data storage

Depends on # of models built and maintained

This is a subjective decision - maybe ideal to have as many models as there are SKU-store combos?

This is probably inefficient: likely some commonalities between similar products or products in a single store.

Here, end decision was 1 model for each product category (e.g., "Bakery") for all stores

## 3. B) Expected annual labor cost for Eurogrocer team

Note there are ongoing **fixed** costs and ongoing **variable** costs \* These are inferred from the pilot data

Model Training	Model Running	Data Storage
15,000 SKU-store combos (pilot)	15,000 SKU-store combos (pilot)	15,000 SKU-store combos (pilot)
16 hours training time (pilot)	4 hours run time (pilot)	95 GB storage (pilot)
\$7.894 compute cost / hour	\$7.894 compute cost / hour	\$.01 GB / month
27 trainings / year	365 runs / year	
= \$7.894 * 27* 16 / 15,000	= \$7.894 * 365* 4 / 15,000 = \$0.77 to run 1 model for 1	= \$0.01 * 12 mos * 95 / 15,000 = \$0.00076 to store data for 1
= \$0.23 to train 1 model for 1 SKU-store combo in 1 year	SKU-store combo in 1 year	model for 1 SKU-store combo in 1 year

## 3. B) Expected annual labor cost for Eurogrocer team

Note there are ongoing **fixed** costs and ongoing **variable** costs

About \$1 / 1 SKU-store combo per year.

Model Training	Model Running	Data Storage	
15,000 SKU-store combos (pilot)	15,000 SKU-store combos (pilot)	15,000 SKU-store combos (pilot)	
16 hours training time (pilot)	4 hours run time (pilot)	95 GB storage (pilot)	
\$7.894 compute cost / hour	\$7.894 compute cost / hour	\$.01 GB / month	
27 trainings / year	365 runs / year		
= \$7.894 * 27* 16 / 15,000	= \$7.894 * 365* 4 / 15,000	= \$0.01 * 12 mos * 95 / 15,000	
= \$0.23 to train 1 model for 1 SKU-store combo in 1 year	= \$0.77 to run 1 model for 1 SKU-store combo in 1 year	= \$0.00076 to store data for 1 model for 1 SKU-store combo in 1 year	7

## 3. B) Expected annual labor cost for Eurogrocer team

Note there are ongoing **fixed** costs and ongoing **variable** costs

Model Training	Model Running	Data Storage
15,000 SKU-store combos (pilot)	15,000 SKU-store combos (pilot)	15,000 SKU-store combos (pilot)
16 hours training time (pilot)	4 hours run time (pilot)	95 GB storage (pilot)
\$7.894 compute cost / hour	\$7.894 compute cost / hour	\$.01 GB / month
27 trainings / year	365 runs / year	

See Excel spreadsheet for calculations by category

# Q4. NET BENEFIT = (BENEFITS - COSTS)

4. A) Net benefit = benefit - cost

See Excel spreadsheet for calculations by category

Annual benefit (sum over all categories) = \$104.12M

Annual variable cost (same) = \$15.36M

Annual fixed cost = \$242.8k

**Upfront cost over 6 months = \$1.765M** 

Benefits are well worth the costs here.

# Q4. NET BENEFIT = (BENEFITS - COSTS)

4. A) Net benefit = benefit - cost

See Excel spreadsheet for calculations by category

Annual benefit (sum over all categories) = \$104.12M

Annual variable cost (same) = \$15.36M

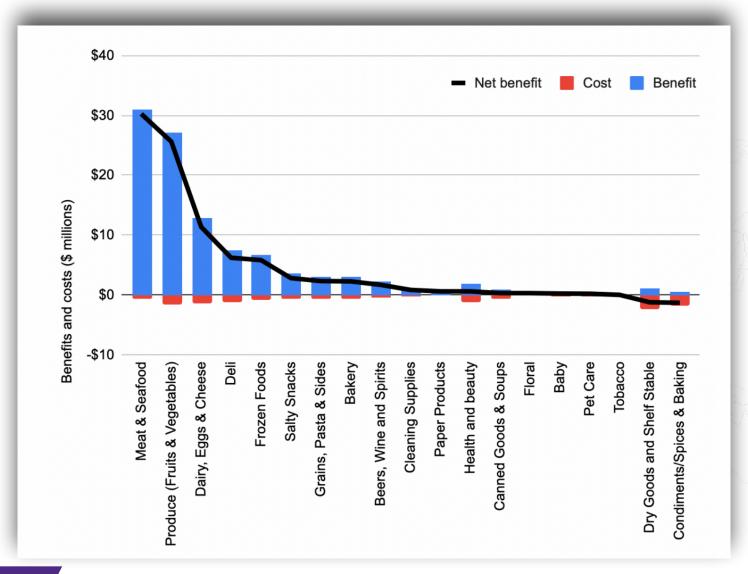
Annual fixed cost = \$242.8k

**Upfront cost over 6 months = \$1.765M** 

Benefits are well worth the costs here.

Which product categories have a negative variable net benefit here?

Why do you think that is?



# OTHER CONSIDERATIONS

#### THE BUSINESS CASE MIGHT NOT BE SO SIMPLE...

The business case for investing in deep learning depends on the upfront investment versus the ongoing benefit. If an immediate payback is needed, that shortens the time frame to realize the benefits, and might affect the decision of whether deep learning is worthwhile.

This doesn't need to be an all or nothing. It's possible that simpler, less expensive, modeling might do a decent job at reducing spoilage and stock-outs.

#### ARE THE BENEFITS / COSTS ACCURATE?

Cost of repeated stock-outs might be > than foregone profit, as customers may switch to competitors.

Reducing spoilage might promote sustainability / corporate social responsibility and prove valuable beyond pure cost of wasted purchases.

# OTHER CONSIDERATIONS

#### **ASSUMPTIONS**

Is food that might spoil truly spoiling, or perhaps being donated at the end of the day (thus yielding different impact of prevent spoilage)?

Are there regulations governing these decisions: e.g., certain level of baby food always needs to be available, or no more than x spoilage.

# THE END RESULT...:(

Despite a strong pilot effort, the roll-out at Eurogrocer was NOT implemented. Why do you think that was?



# THE END RESULT...:(

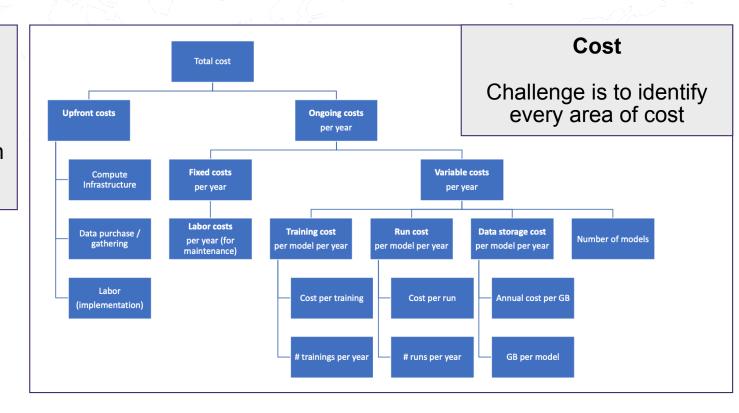
Despite a strong pilot effort, the roll-out at Eurogrocer was NOT implemented:

- 1. Data on store closings and personnel absences were missing from data, which turned out to be really important in the ensuing months.
- 2. Some of the data was wrong: for instance, data might show that some products were on promotion, but the store managers knew that was wrong. That started to erode managers' confidence in following the model recommendations.
- 3. The pilot had prediction error at about 18%, but expectations were 10%. Even though 18% is still profitable, the inability to hit the projected 10% was questioned, and made the full roll-out suspect.
- 4. Cloud costs increased QUICKLY. These costs were mitigated quickly, but the spike made Eurogrocer leaders skeptical.
- 5. Complexity and black-box-ness of deep learning scared away some members of Eurogrocer's board who felt like it was too sophisticated for them. Some board members recognized that maximum predictive accuracy was key, but too many were too skeptical.

This makes you a **better analytics communicator**. Analytics is NOT always worth the cost, and I want you to know how to articulate costs /benefits, so you can make the case for your own work.

#### **Benefit**

Challenge is to estimate:
"How much business value is generated from an increase in predictive accuracy?"



This makes you a **better analytics practitioner**.

You cannot control the benefits of your analytics work. You can build a great model, but the value of the model is determined by the given context, not you.

You CAN control the costs of the work you do, and you CAN make recommendations about the amount of work you should do, given the benefits.

This is a core foundation of **analytics leadership**.

This makes you smarter about how executive decision-making works.

Senior business leaders are constantly concerned with capital allocation. Despite the high ROI, the executive team has other alternatives to use their capital, and this project was judged too sophisticated and high-risk.

It is easy to look at the ROI figures and say "this is an easy decision," but the executive team has limited capital and must make decisions on many factors, not just "is the ROI positive?"

In addition, each business functional leader would likely weigh in on a decision to move forward. What are the implications for the head of IT, HR, operations, etc.?

This makes you a better analytics communicator.

This makes you a better analytics practitioner.

This makes you smarter about how executive decision-making works.