Data-driven decision making

Michael @ Delivery Hero

Data Science Seminar
DESY Zeuthen

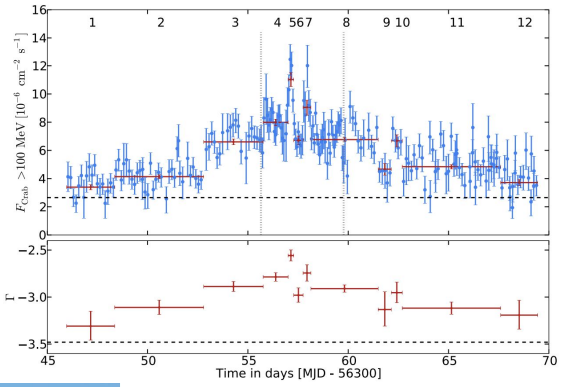
What’s on the Menu

- My Journey from Academia to Industry
- Selected Projects
- Hiring from both sides

My Journey: Academia

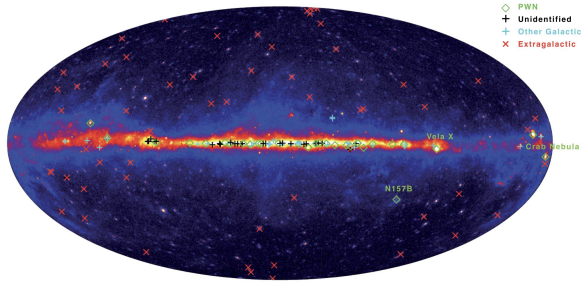
2011 - 2014: PhD@DESY

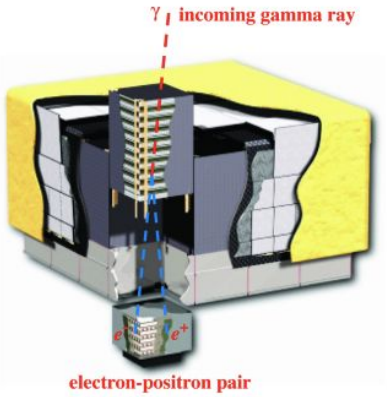
- H.E.S.S.: N157B, PWN modelling
- Fermi-LAT: Crab Nebula Flares
- GammaLib / ctools: PWN population MeV - TeV, combining Fermi and HESS



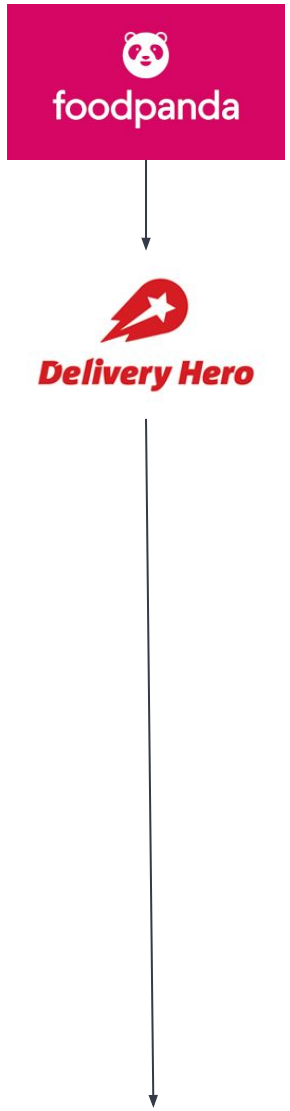
2014-2017: PostDoc @ HU

- CTA, MST Prototype
- GamaLib / ctools development
- H.E.S.S. analyses





My Journey: Industry



Feb 15th 2017, 9:00-9:20am: Data Scientist @ Foodpanda

Feb 15h 2017 9:30am: Data Scientist @ Delivery Hero

- Team size 3
- Working on data science projects

June 2018: Senior Data Scientist

- Team size: 6 (1 Head, 2 Senior DS, 2 Data Scientist, 1 working student)
- “Owning” data science projects and guide more junior team members
- Take care of stakeholder management

May 2019: Head of Data Science

- Dumb Luck: My boss left and I got offered to be his successor
- Team size: 10 (2 Seniors, 1 Engineer, 5 Data Scientists, 1 working student and I)
- Meetings, meetings, meetings, managing stakeholder relationships, saying “no” to incoming requests
- Managing and guiding people, distribute work, inner-team hierarchies and processes

About Delivery Hero

Hard Facts

- Founded 2011 in Sweden
- Operating food delivery companies in 40+ countries with 22,000 employees (~1,500 in Berlin HQ)
- Largest food network outside China (500,000+ restaurants on our platforms)
- Currently processing > 3M orders per day (YoY growth ~100%)
- Offering own delivery services in 500+ cities
- Public company: 15bn EUR valutation
- Not operating in Germany anymore (sold Lieferheld, Pizza.de and Foodora to our competitor Takeaway.com)
- Currently heavily investing in “Q-commerce”

DH Headquarters, next to Monbijoupark in Mitte

My personal mission

Challenge gut-feelings in business decisions by building data products that enable data-driven decisions in a business context.

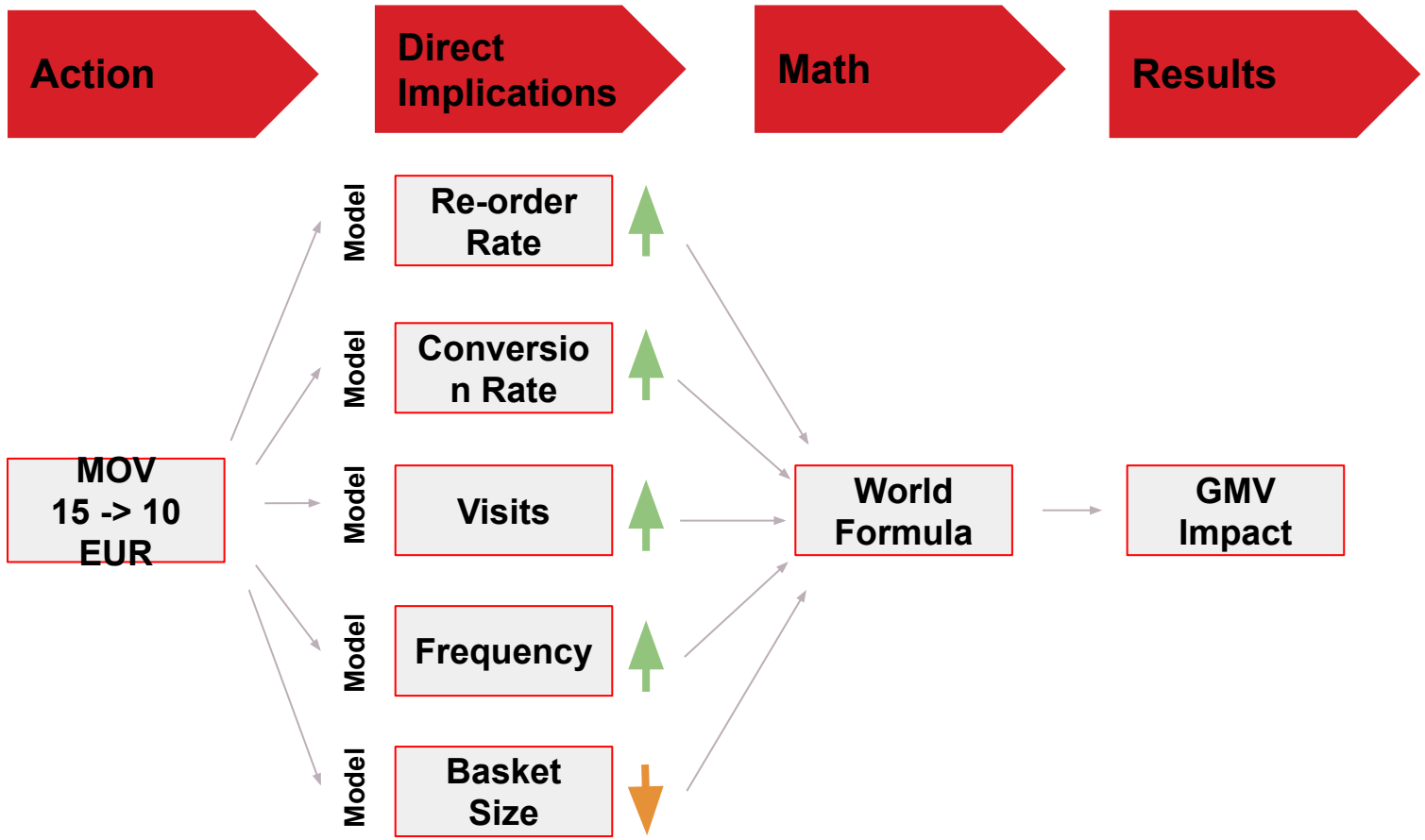
Selected Projects I

… and where physics helps

Predict Impact of Price Changes

There are two price points we can negotiate with the restaurants: minimum order value (MOV) and delivery fee (DF).

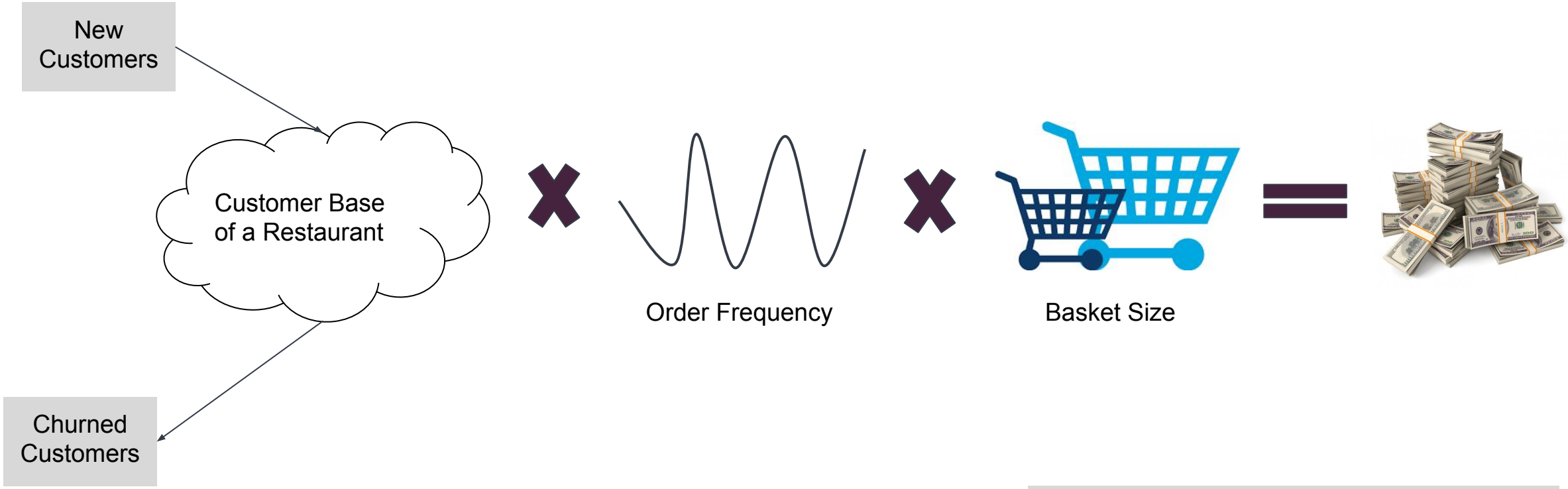
Business Question: How will a change in MOV and DF affect the restaurants' business (e.g. revenue)



Predict Impact of Price Changes

There are two price points we can negotiate with the restaurants: minimum order value (MOV) and delivery fee (DF).

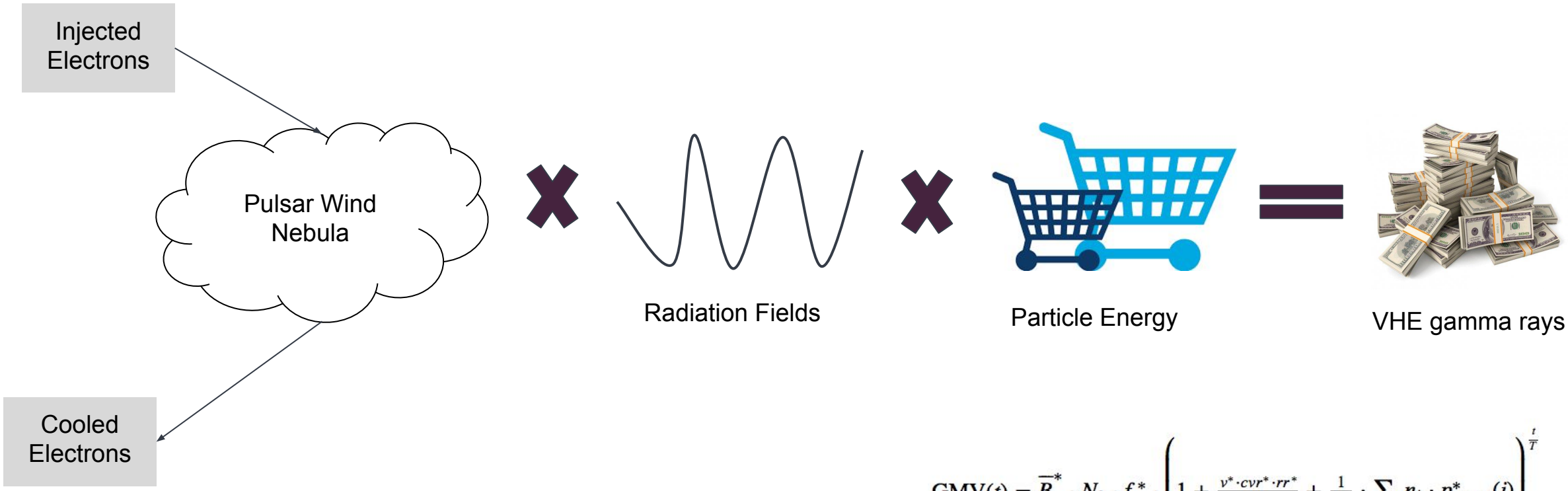
Business Question: How will a change in MOV and DF affect the restaurants' business (e.g. revenue)



Pricing affects all of these points

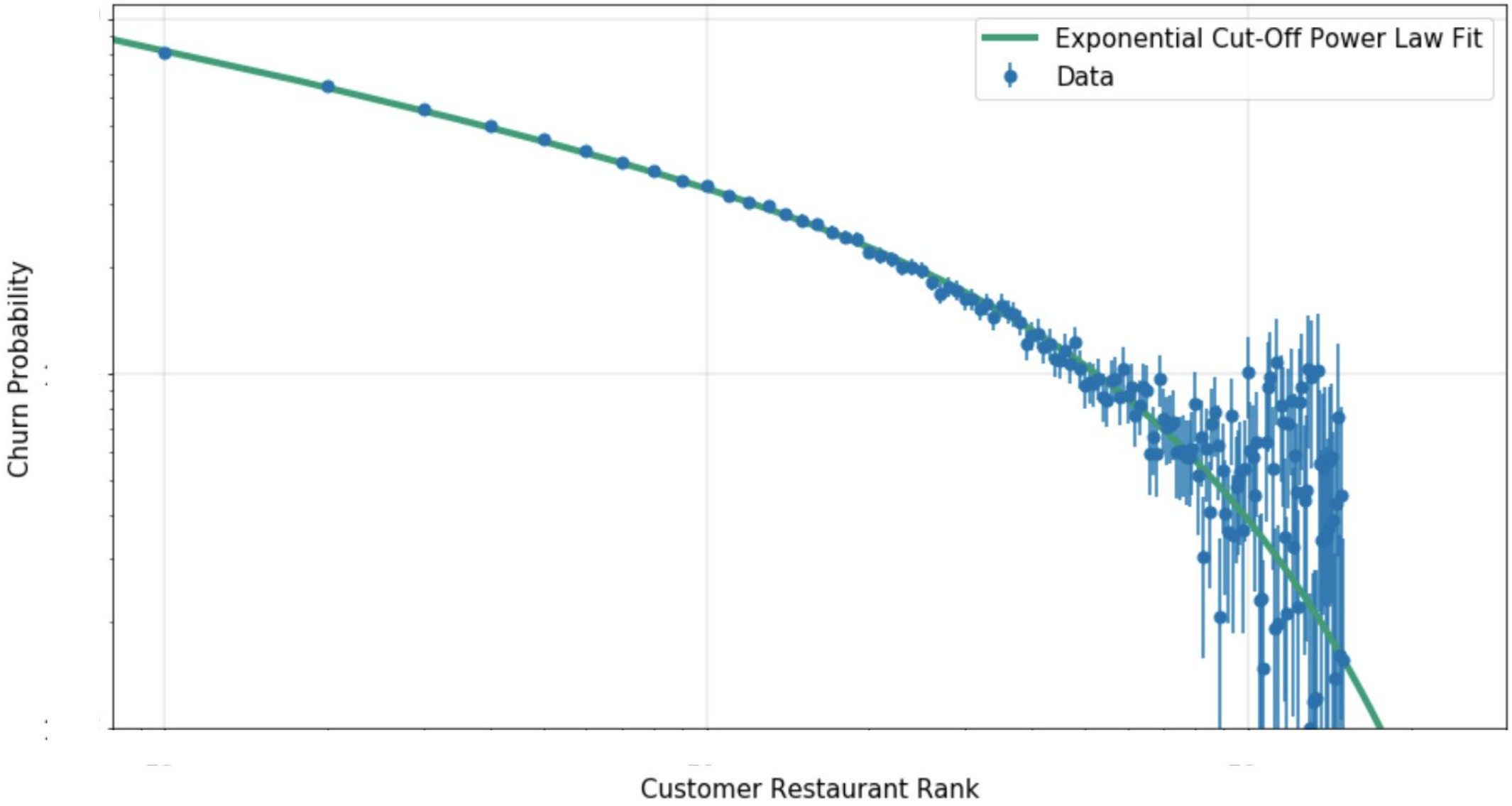
Analogy to Astrophysics

This problem sounded too familiar ;)



$$\mathrm{GMV}(t) = \overline{B}^{*} \cdot N_0 \cdot f^{*} \cdot \left(1 + \underbrace{\frac{v^{*} \cdot cvr^{*} \cdot rr^{*}}{N_0}}_{\text{New}} + \underbrace{\frac{1}{N_0} \cdot \sum_i n_i \cdot p^{*}_{\text{churn}}(i)}_{\text{Churned}} \right)^{\frac{t}{T}}$$

Example: Churn Probability



Selected Projects II

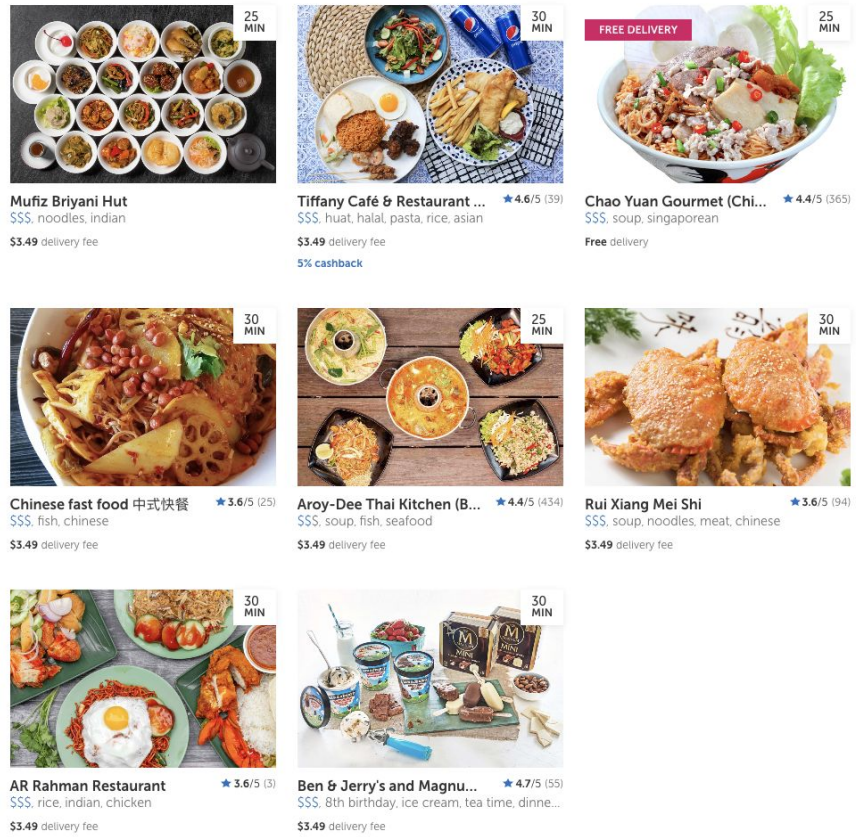
Optimising Choice - Geospatial Analyses

Business Questions

- How many restaurants should be delivering to a specific customer location?
- What's the ideal split between cheap, expensive, Italian, Sushi...?
- How many more orders could we generate by optimising the supply in a specific location

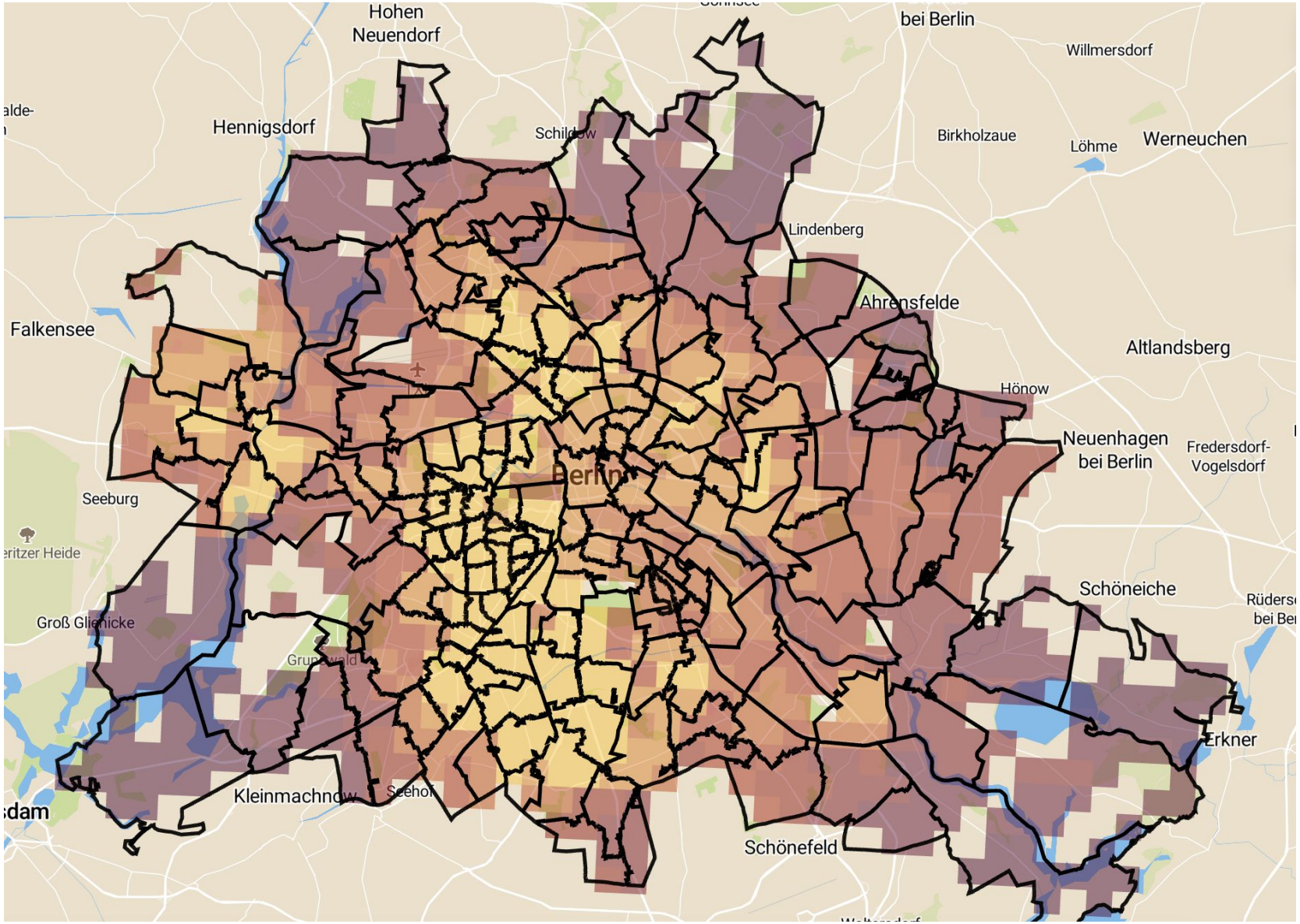
Implied Actions

- Bottomline: Data-driven decision making which restaurant to onboard where in order to fill a gap



Still the best tools: heatmaps

- Define Inventory Goodness
- Develop self-service data product to play with color scales, KPIs etc
- Business owners (Sales) would plan and prioritize their efforts based on this tool
- Enabling and advertising data-driven decision making



Hiring

Sitting on both sides of the table

My application anekdotes

Coding Challenges

- For a job as “Risk Data Scientist”, I was asked to “design the software architecture of a risk detection system to detect fraud” and brainstorm with the person via email “for a couple of weeks”.
- For a travel company, I got sent a 5.3GB zipped (!) csv file of flight data with the vague instruction: “tell us your findings” (the file didn’t even fit into memory)
- Sent a perfect solution to a coding challenge including a bonus question only to get a standardised rejection 1min after I handed in the solution

My application anekdotes

Hiring Process

- After 5 stages(!) of interview and a coding challenge I was invited to a “trial day”. The latter got cancelled one day before because they made a “mistake in the calculation of headcount and the position is now closed”
- Had to prepare a 45 min presentation about myself for the whole team, only they could tell me afterwards they were actually looking for a software engineer and therefore went with someone else
- One guy kept on talking how “his other data scientist” is doing weird things and he wants a second eye on them…

→ **Brace for turbulences and idiotic things to happen**

My application anekdotes

My current Job

- via recruiter
- no coding challenge
- always quick responses
- two lovely conversations with my boss and then with his boss

→ Trust your instincts!


Switching sides

Switching sides

Note these numbers are approximate (for 7 people I hired)

- scanned through ~600 CVs
- ~50 interviews (mostly via phone)
- Gave people a chance I shouldn't have, rejected people I shouldn't have → learning curve
- Have a lot of help of our Department for Talent Acquisition in screening CVs and taking initial calls

The perfect candidate from a Hiring Manager point of view

- Technical skills are given (coding / statistics / machine learning / databases / tools)
- Knows how to explain results and processes to non-technical audience (i.e. stakeholders)
- Can work independently, creatively solves any given problem and **doesn't need much input and guidance**

So how to find out these things during an interview?

Switching sides

What’s the difference between the mean and the median?

A1: One is the sum of the values divided by the length, while the other is the middle value of an array

A2: there are 10 people in a bar, everyone has a salary of 50,000 EUR per year, then Lionel Messi comes in. The Median of salaries would stay the same because it is much more robust to outliers, while the mean would be changed to somewhere above 1M EUR. So the mean can depend very much on one outlier.

+ Give examples and explain creatively

Switching sides

How do you access the last element of a list in Python?

A1: brackets -1

+ Simple question, simple answer, don’t talk more than needed in such cases

Switching sides

What ways of target setting and goal setting are you used to?

A1: I usually come up with topics I wanna work on and pitch them to my supervisors, then in close alignment with the company strategy, we build a plan how we can support the business best

A2: Someone usually tells me what to do

+ Show that you're proactive and enthusiastic for enriching the business independently

Summary hiring

Things I'd recommend for an interview

- Never take it personal!
- Salary question: know what you're worth and say it, don't explain your salary expectation
- If not mentioned by the interviewer, ask about the team size, the team structure and dynamics (shows you're interested in having a good working atmosphere)
- Research the company, their values, and how the business works (I always ask: how do we make money?)

There is light at the end of the tunnel

- Getting the first job is the hardest
- Now I am getting ~2-3 job offers per week via LinkedIn (admittedly many of which is nonsense)

Delivery Hero

Danke for listening

www.deliveryhero.com