

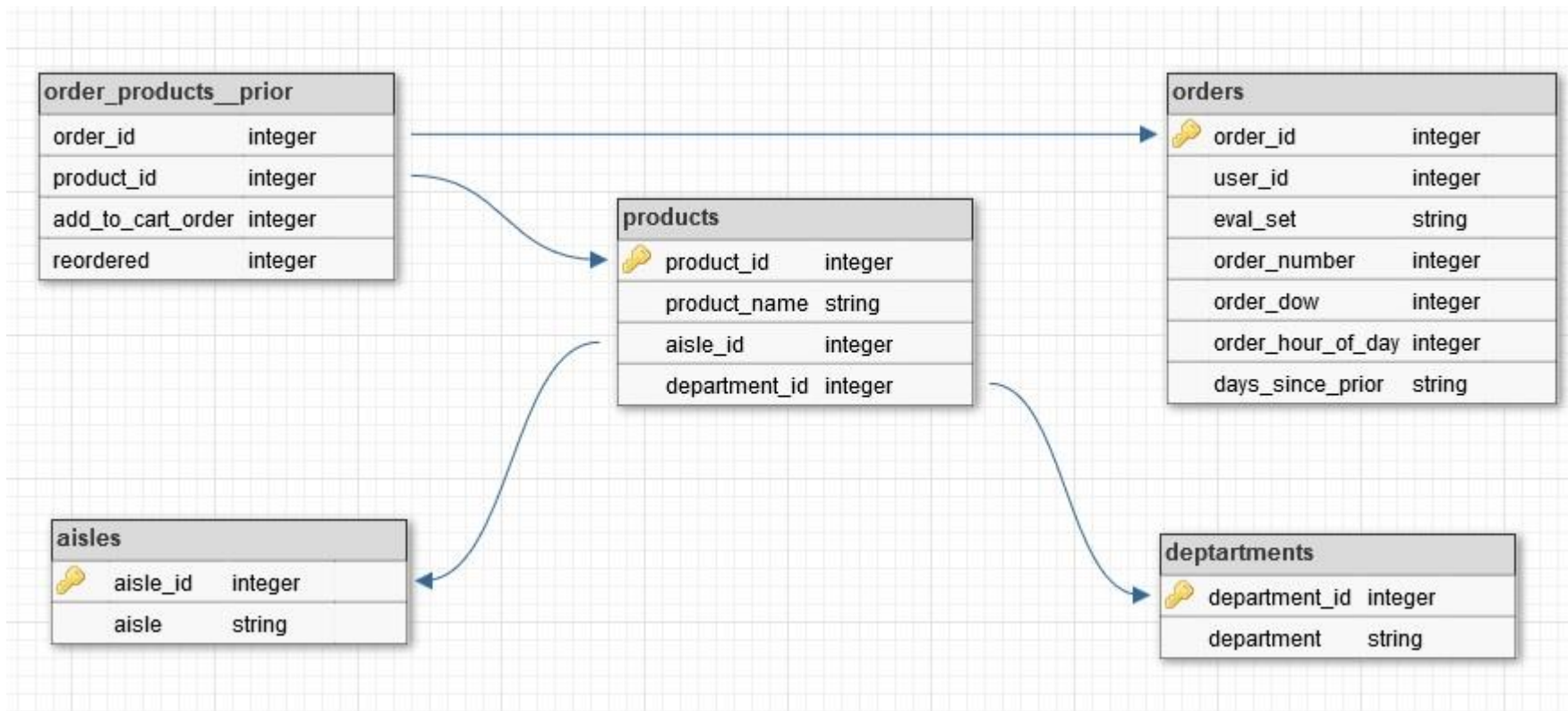


von & Igor Arkhipov &
Jannik Bruhns



Das US-Startup Instacart heuert Privatpersonen an,
um von Nutzern online bestellte Waren
im Supermarkt einzukaufen und anschließend auszuliefern

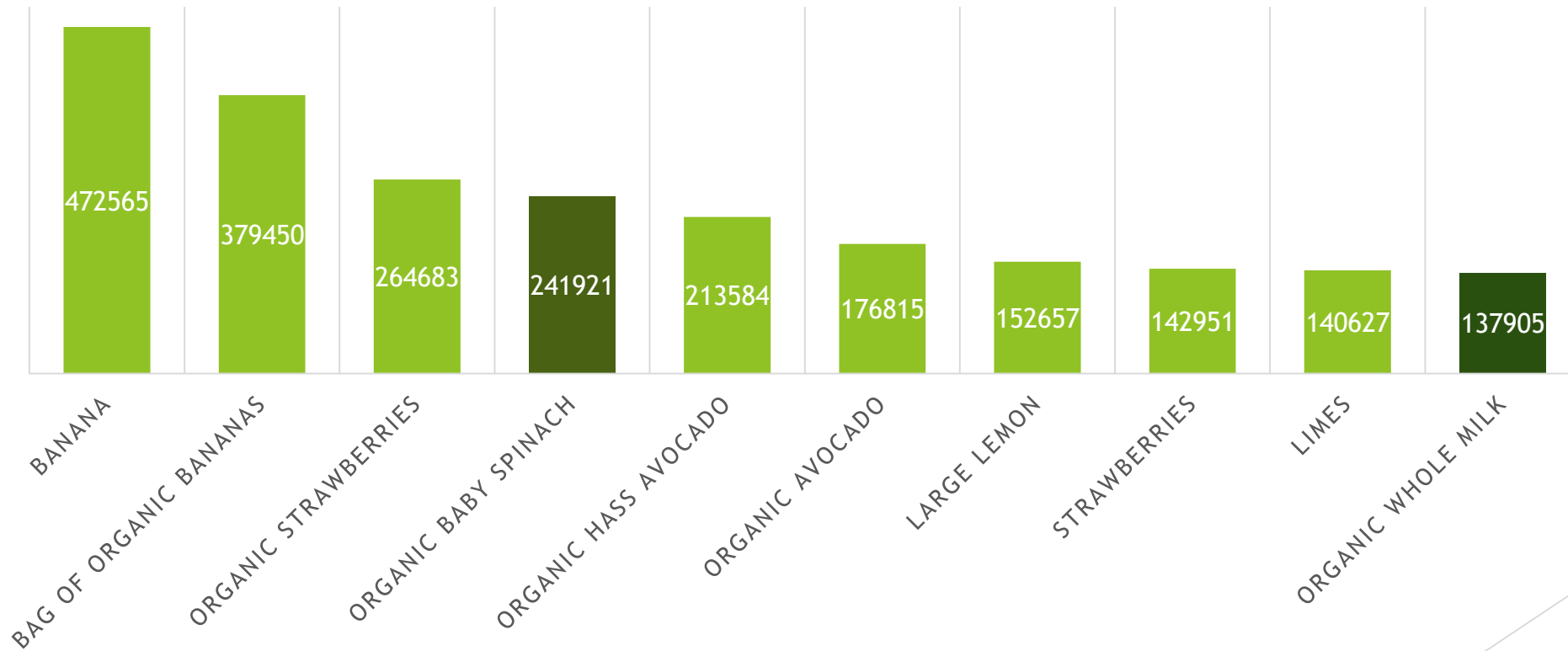
Datenbankdiagramm



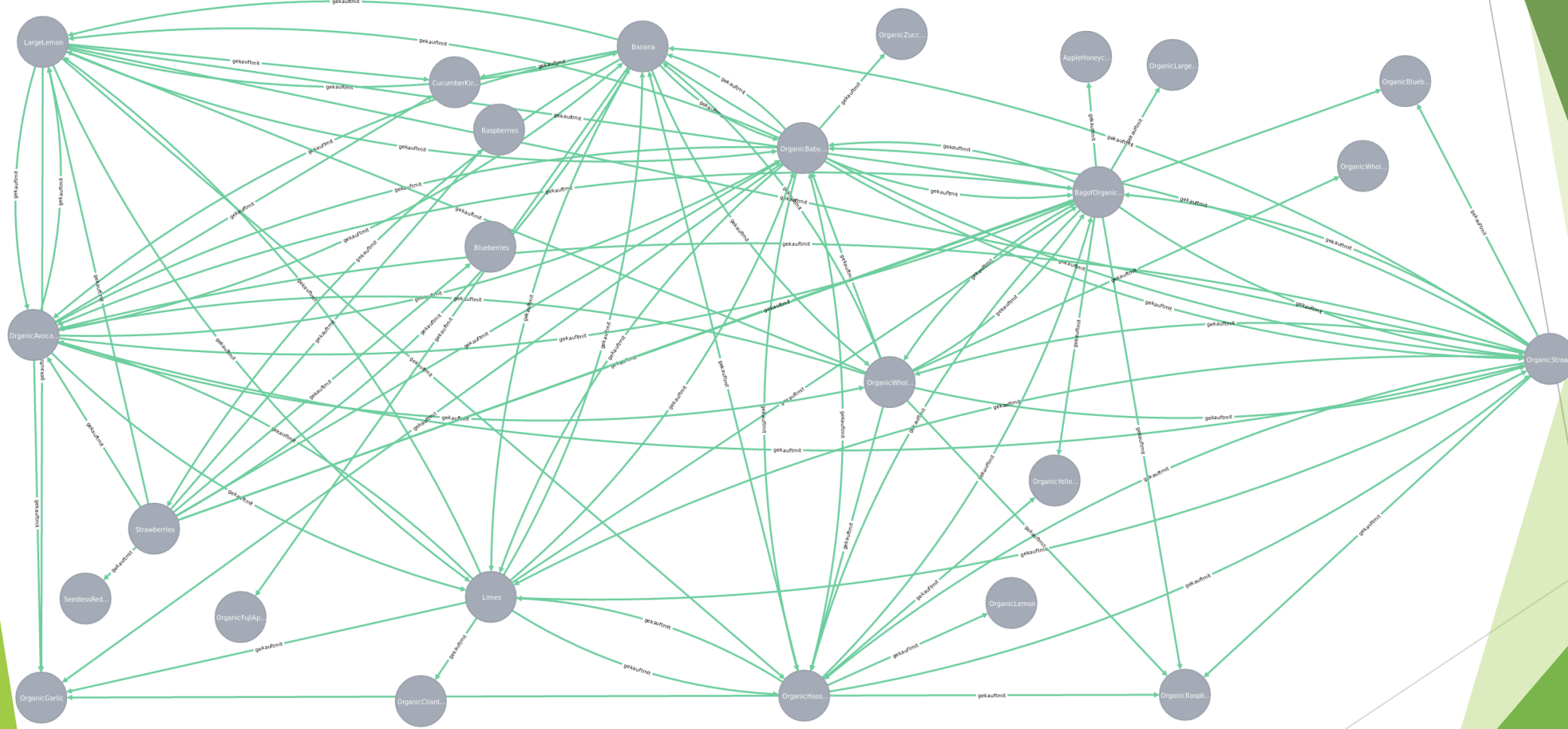
Daten

- ▶ CSV Dateien von der Instacart Webseite
- ▶ Über 3 million orders von mehr als 200,000 Instacart Usern
- ▶ 49.688 Produkte
- ▶ 32.434.489 Bestellpositionen
- ▶ Import ~2 Stunden
- ▶ MongoDB (Dokumentenorientierte Datenbank)
- ▶ Neo4j (Graphdatenbank)

Top 10 gekaufte Produkte

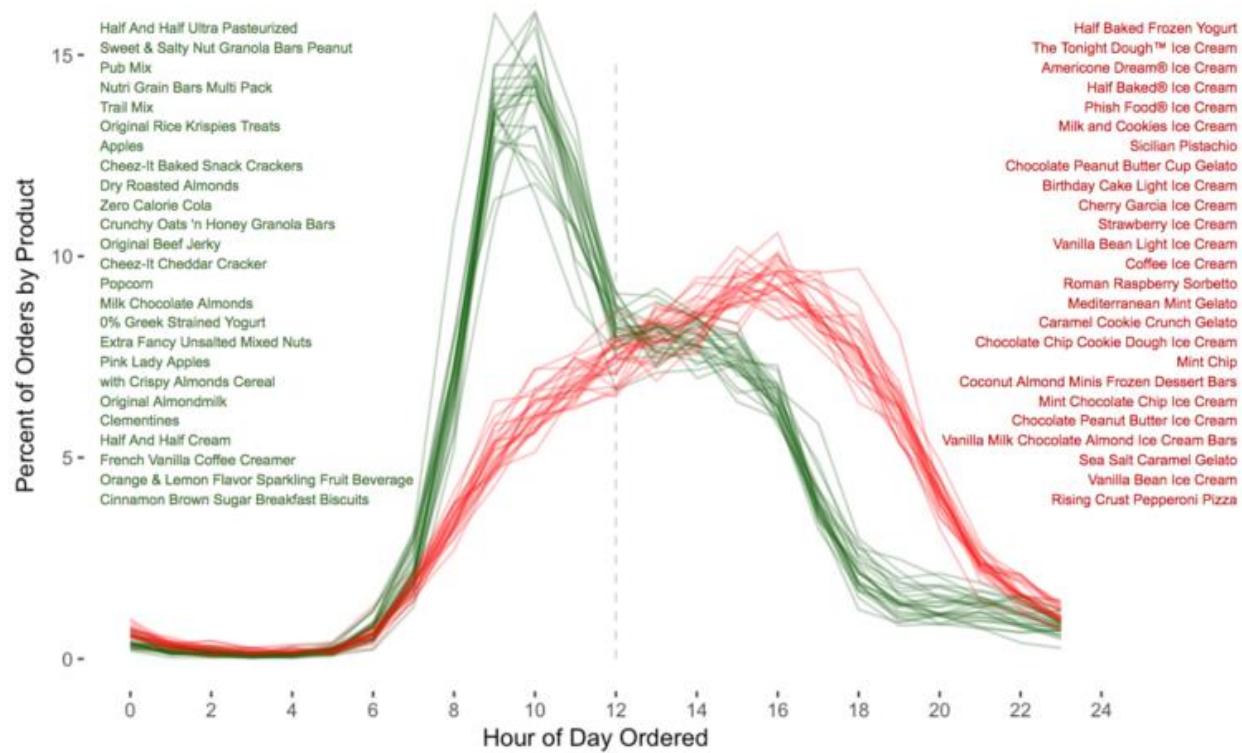


Top 10 Produkte und deren meist mitgekauften Produkte



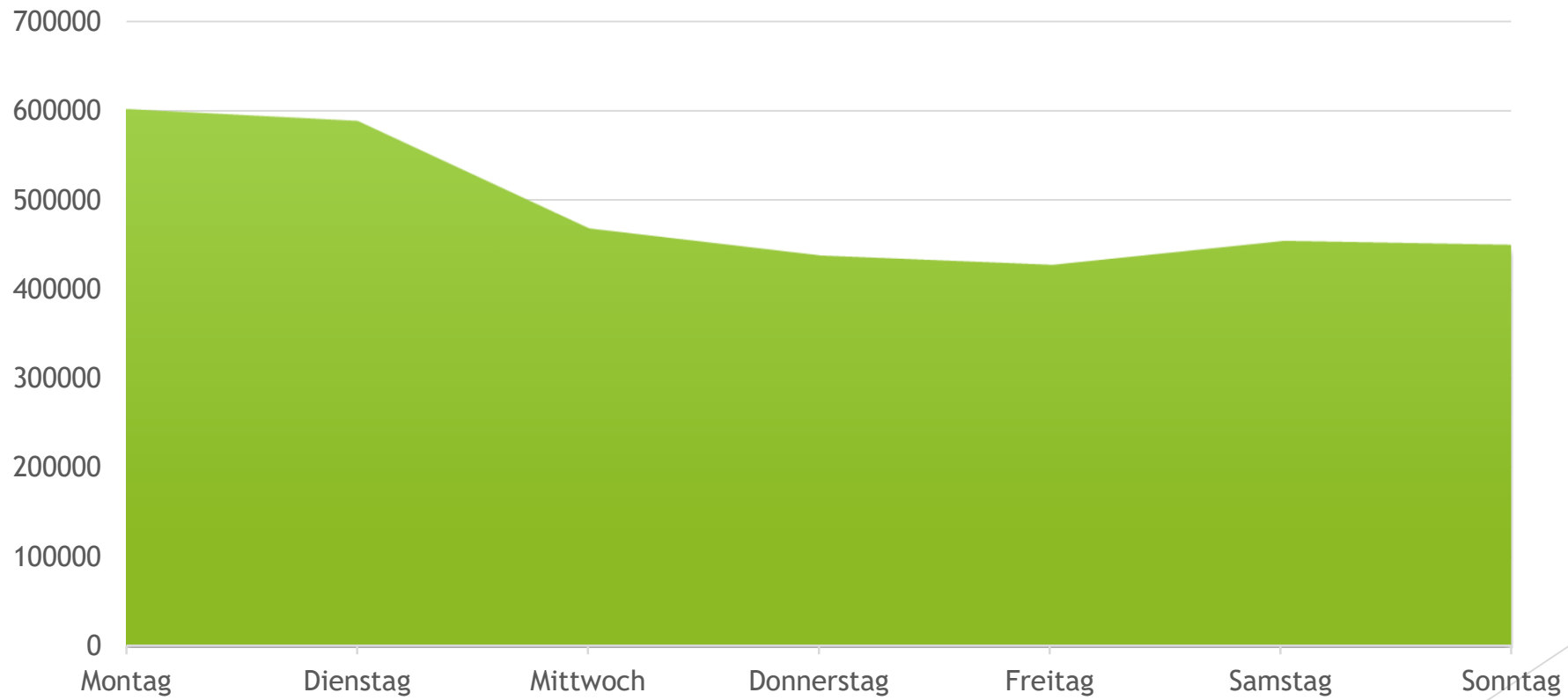


- Milch reorder > Backzutaten

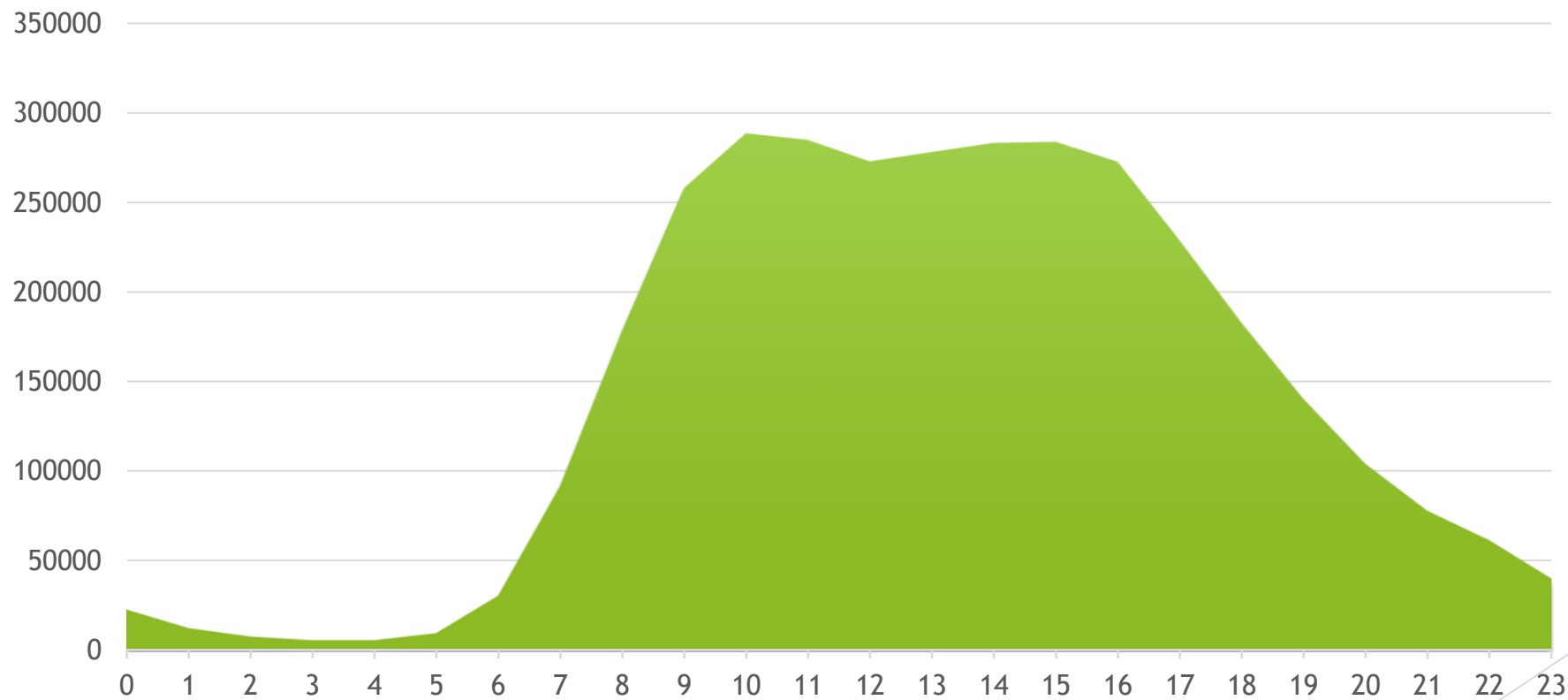


- ▶ Morgens mehr gesundes Essen
- ▶ Abends eher ungesundes Essen
- ▶ Zuletzt gekaufte Produkte Top24 Eis Cream, #25 Tiefkühlpizza

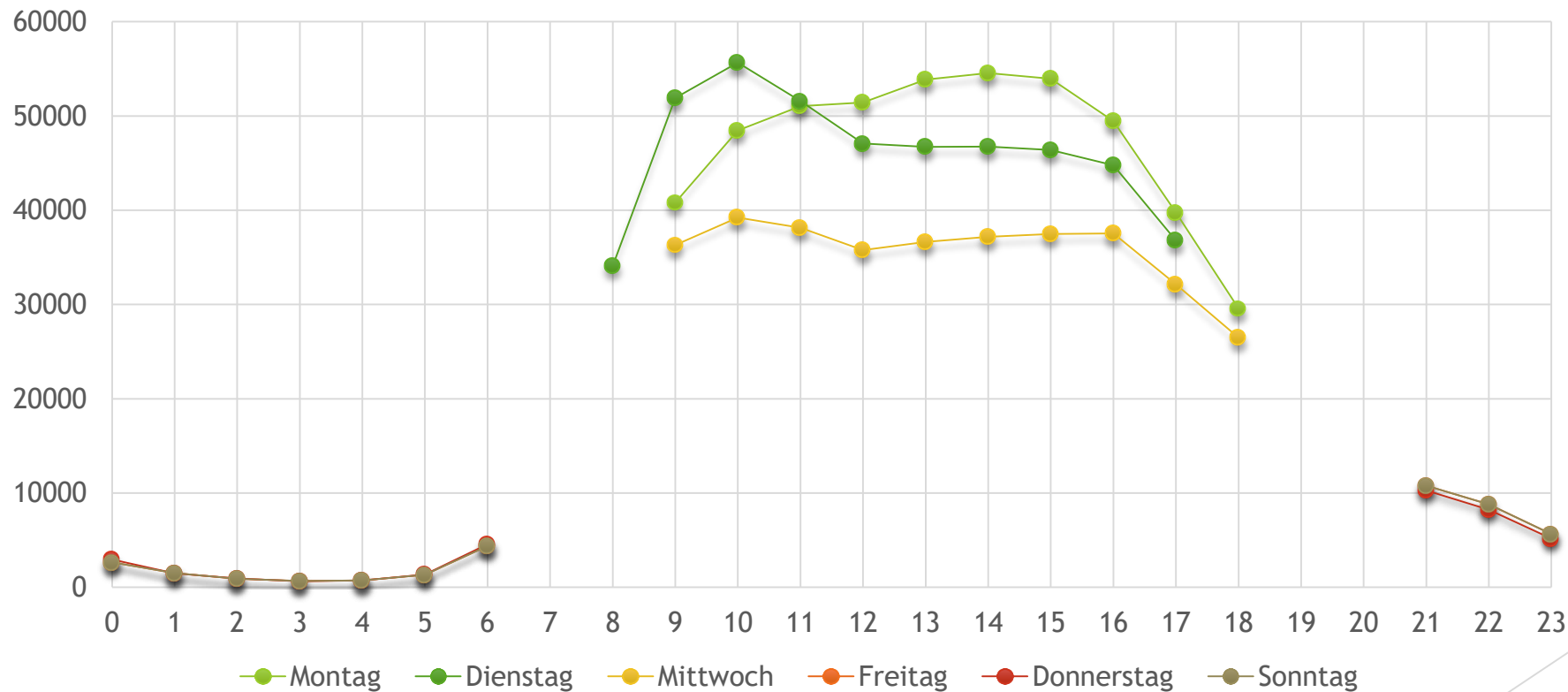
Bestellungen pro Tag



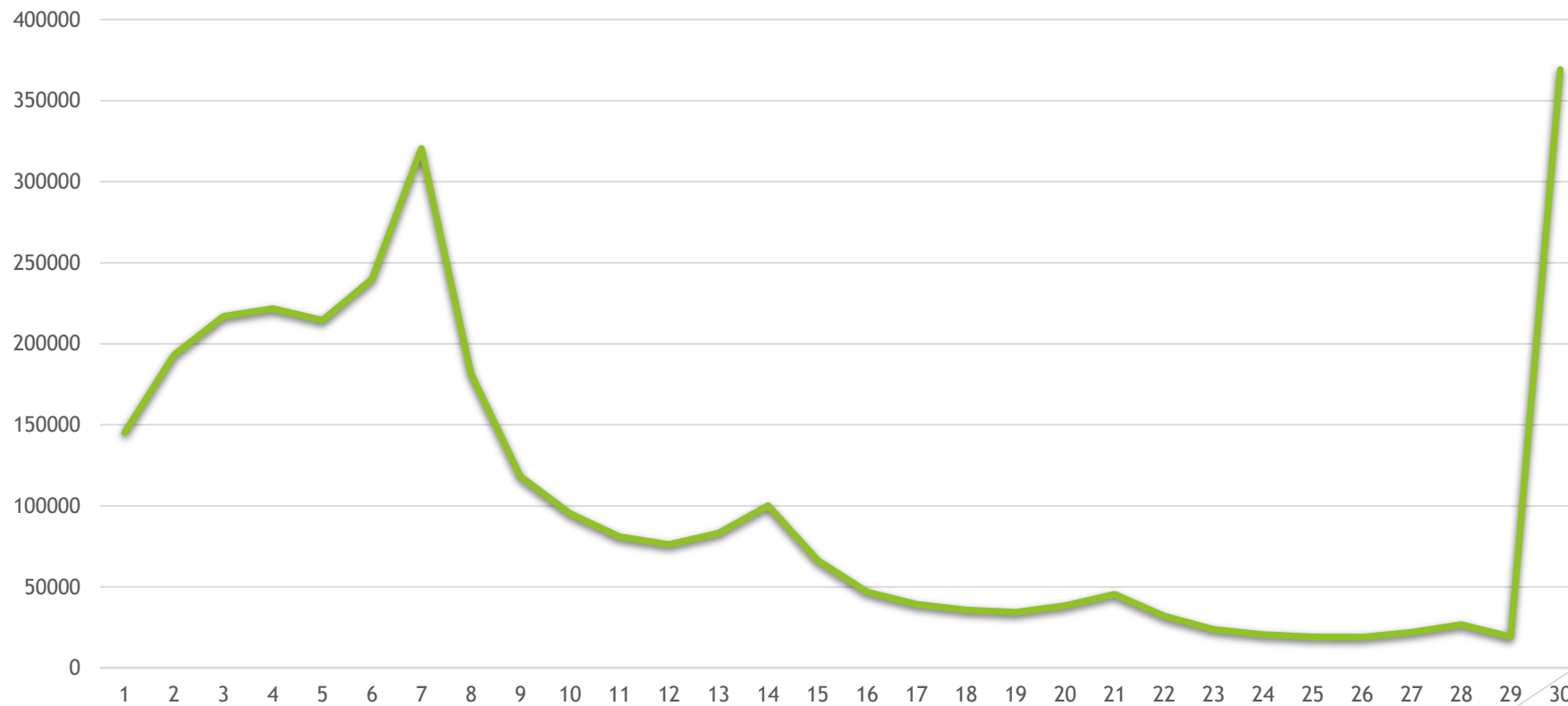
Anzahl Bestellung pro Uhrzeit



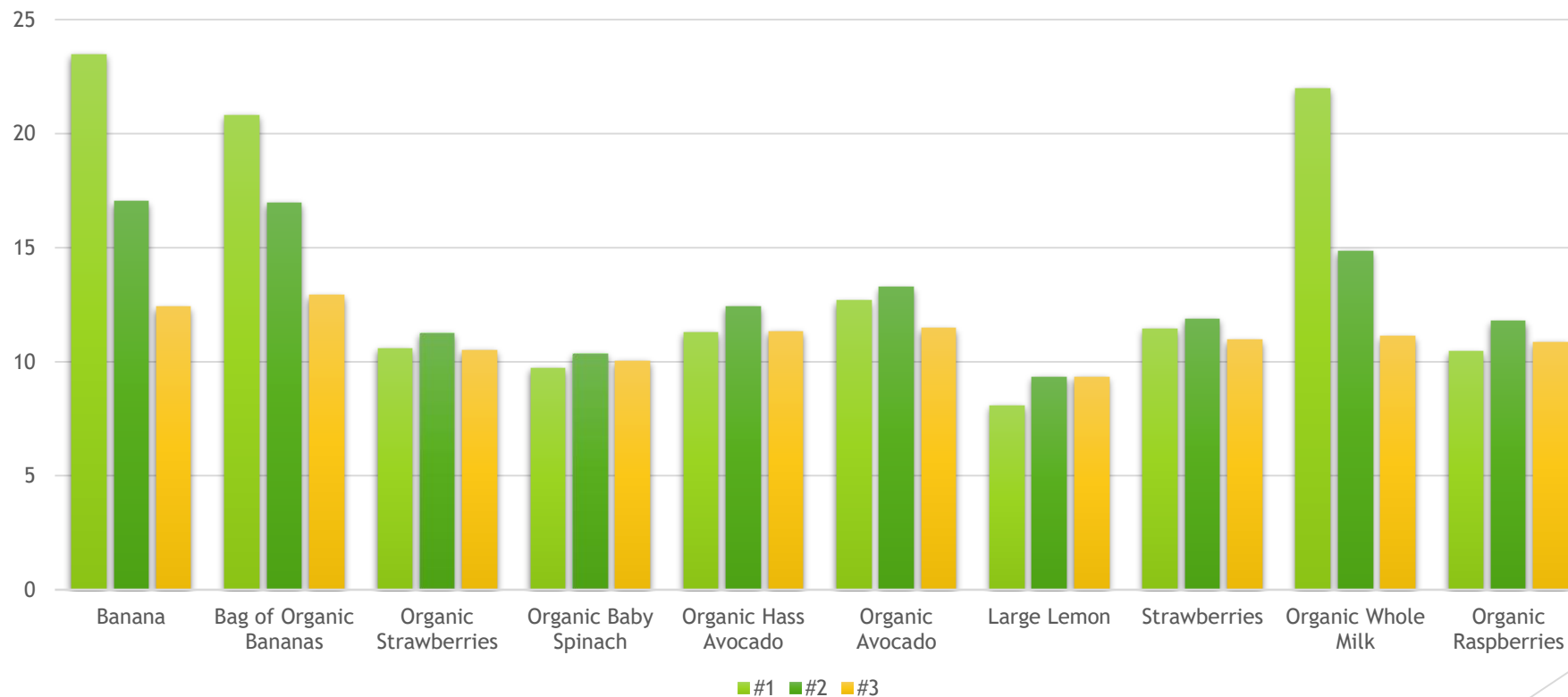
Top 3 Tage und Uhrzeiten + Bottom 3 Tage und Uhrzeiten



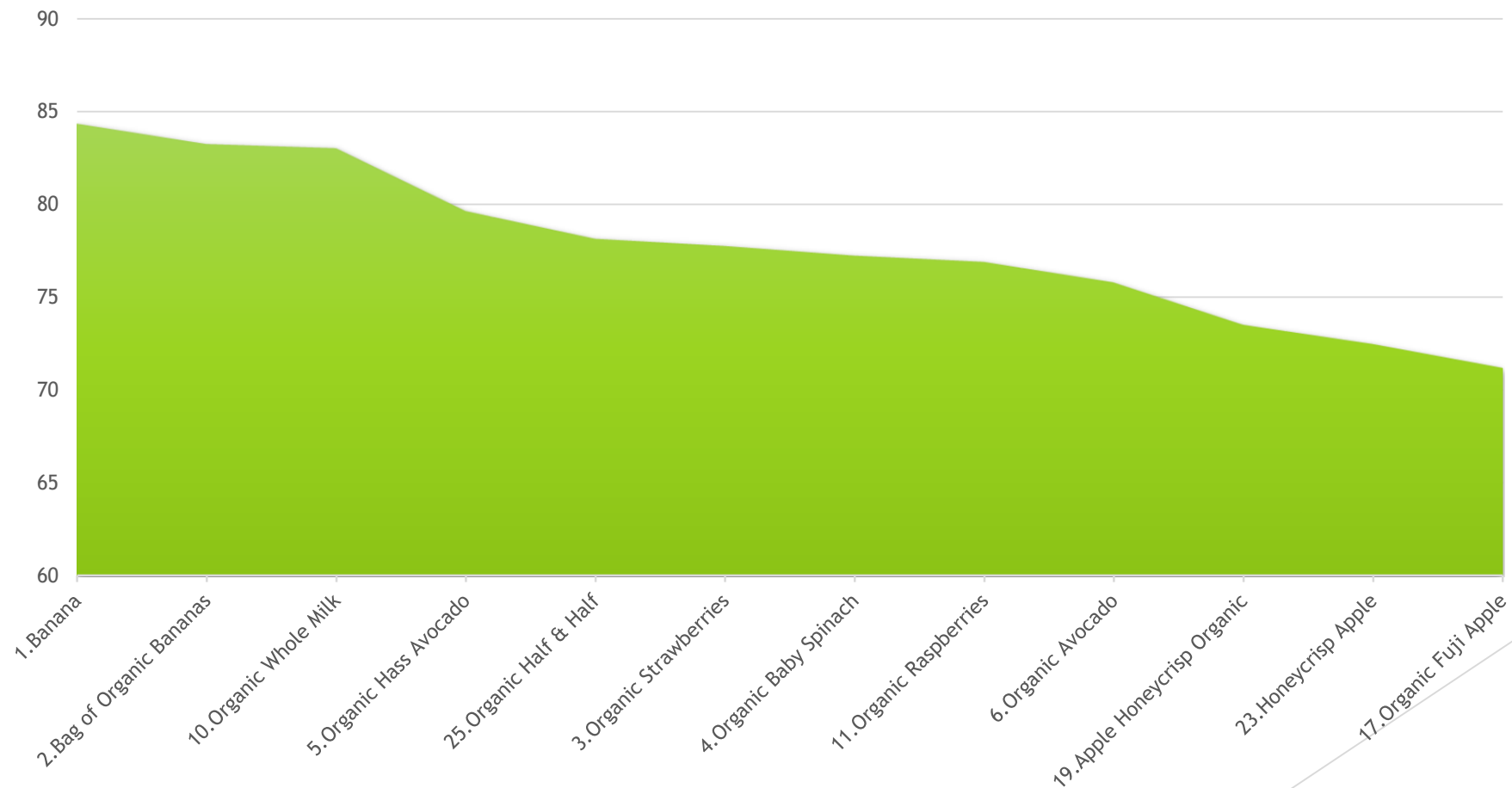
Nach wie vielen tagen die meisten wieder eingekauft haben



Welche Produkte wie oft als erstes, zweites oder drittes in Warenkorb gelegt



Top10 Produkte die wieder bestellt wurden (Prozent)



Probleme

- ▶ Durchschnitt nach wie vielen tagen wieder bestellt wurde
String kein avg(Durchschnitt) Query möglich
- ▶ Top 10 aisles oder departments
 - 10.000 = 24m
 - 100.000 = 240m
 - 1.000.000 = 2400m
 - 10.000.000 = 24000m
 - 32.000.000 = 76800m = 53,3 TAGE