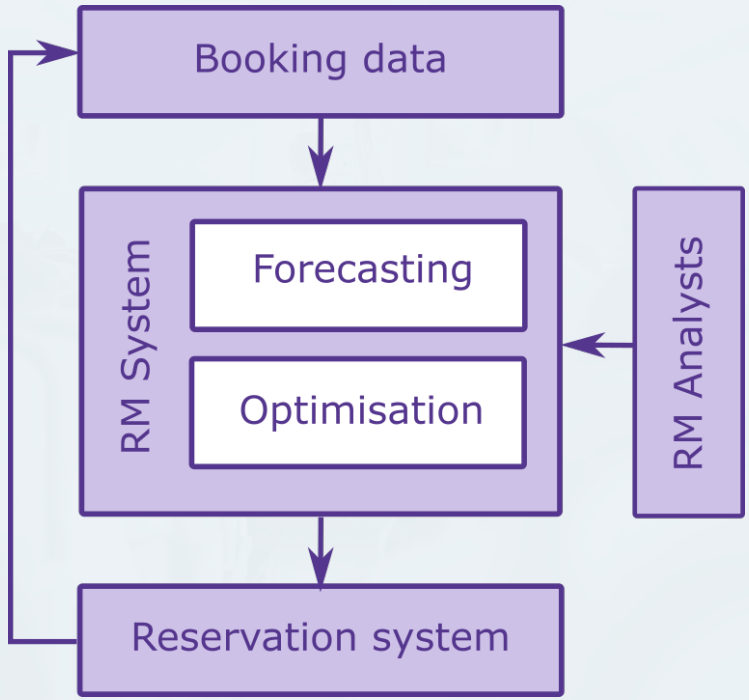
Detecting outlier demand in revenue management networks

3rd IMA & OR Society Conference

Nicola Rennie • Lancaster University

- *Controls ticket availability.*
- *Combines forecasting with optimisation.*
- *Decision support for RM analysts is needed.*



- *Over 1000 stations.*
- *Over 110,000 possible O-Ds.*
- *Bookings are reported on the leg-level.*

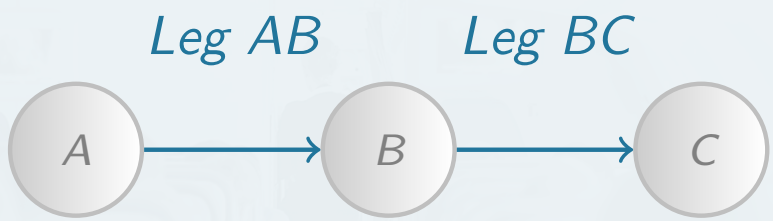


Figure: Railway network with two legs

- *Outliers don't affect entire network, nor single leg.*
- *Partition network using clustering.*

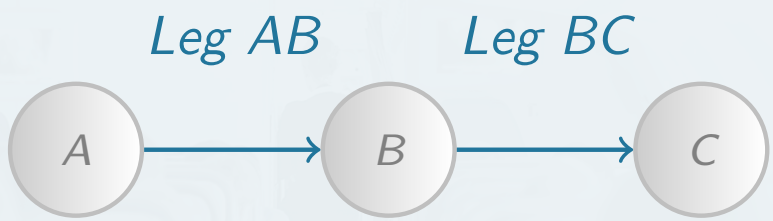


Figure: Railway network with two legs

- *Nodes represent stations.*
- *Edges represent legs connecting stations.*

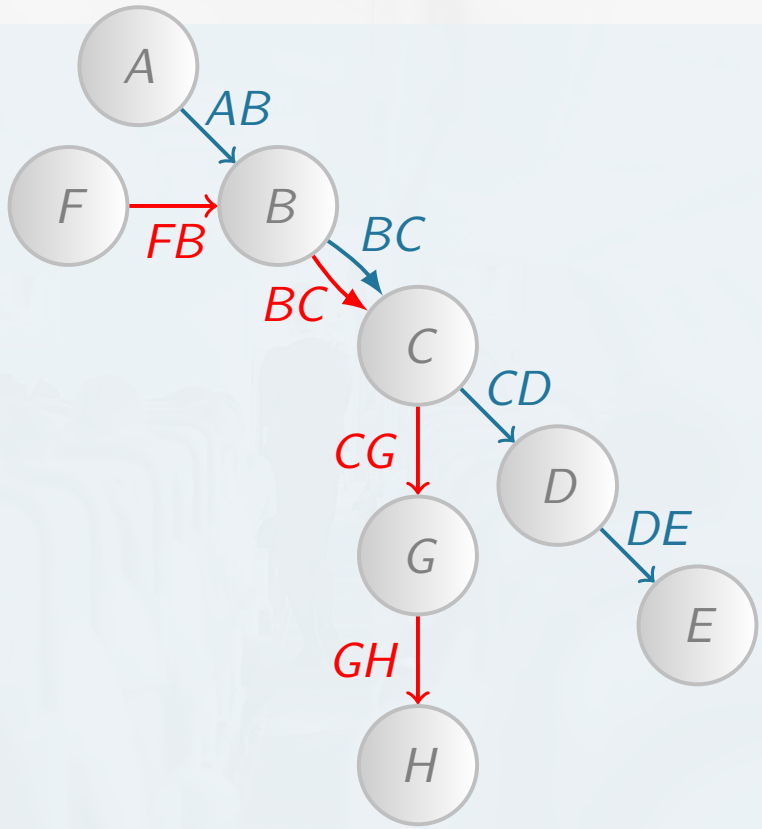


Figure: Railway network graph

- *Nodes represent legs.*
- *Edges define which legs can be in same cluster.*

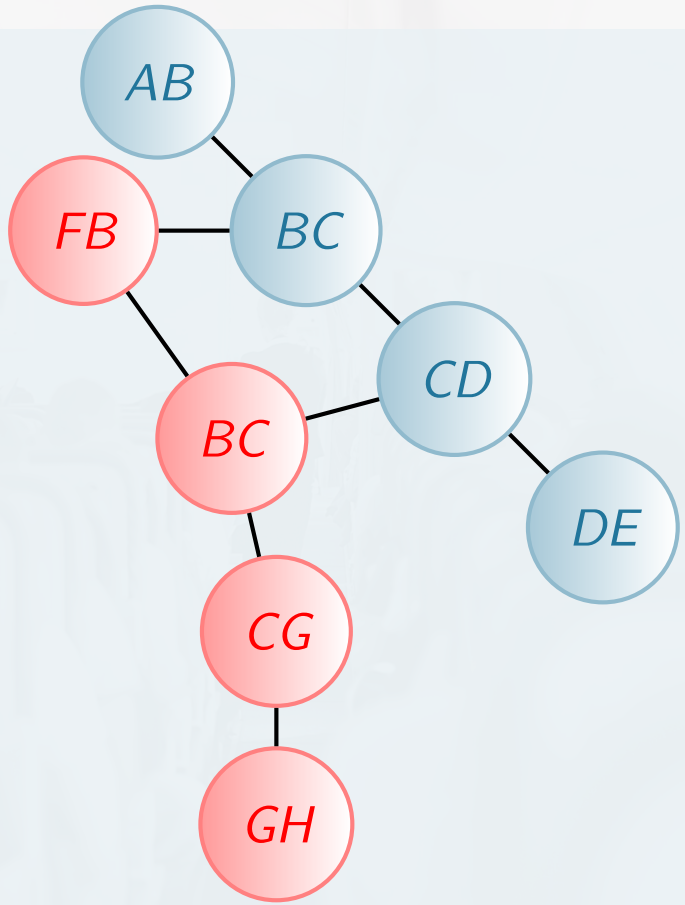


Figure: Inverted graph

- *The common traffic ratio of legs AB and BC is:*

$$r(AB, BC) = \frac{D_{AC}}{D_{AB} + D_{BC} + D_{AC}}.$$

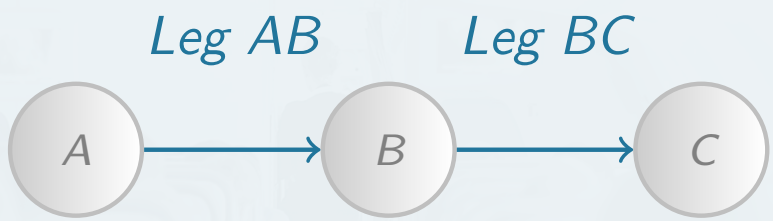


Figure: Railway network with two legs

- *Add in common traffic ratio to edges*

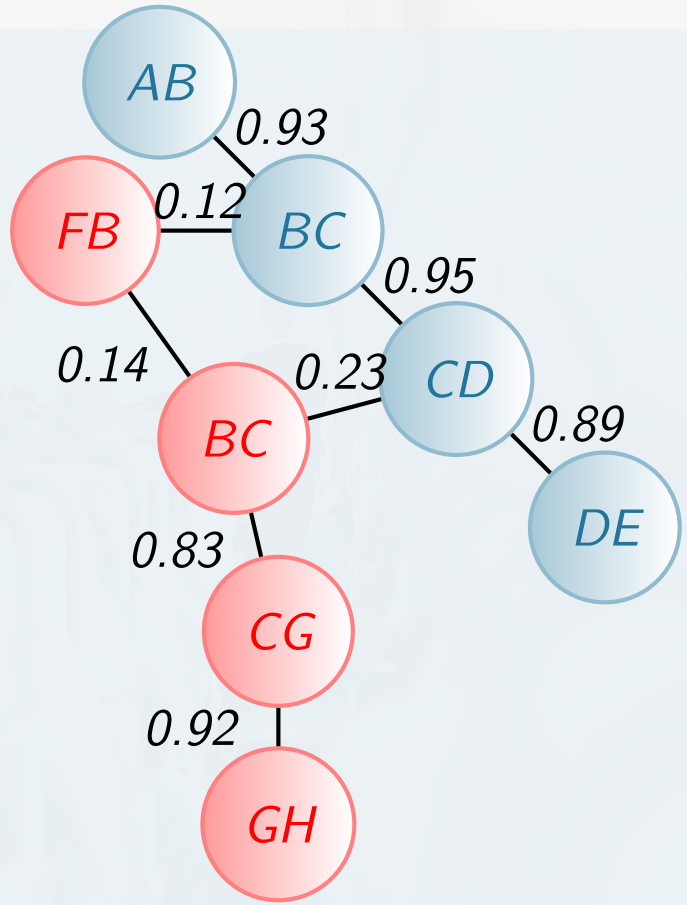


Figure: Inverted graph with common traffic ratios

- *Edge weights are* $1 - r()$
- *Obtain minimum spanning tree (Prim's algorithm)*

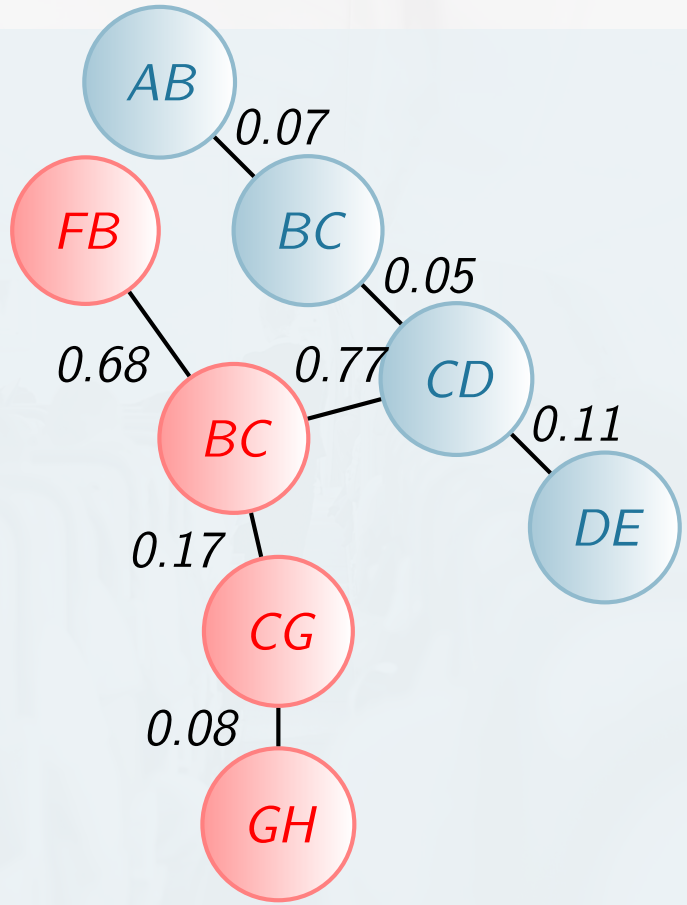


Figure: Minimum spanning tree with edge weights

- *Remove edges with weight above some threshold.*

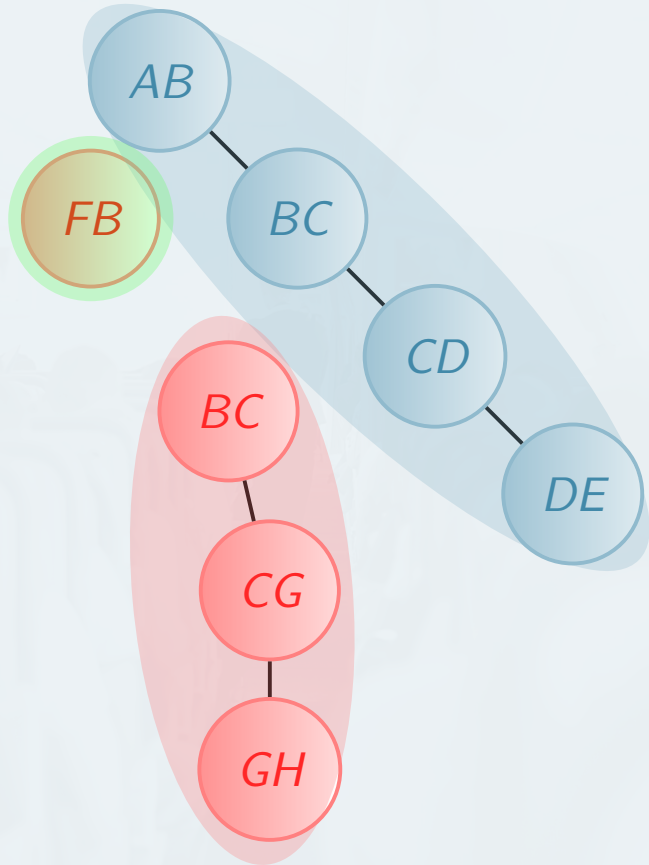
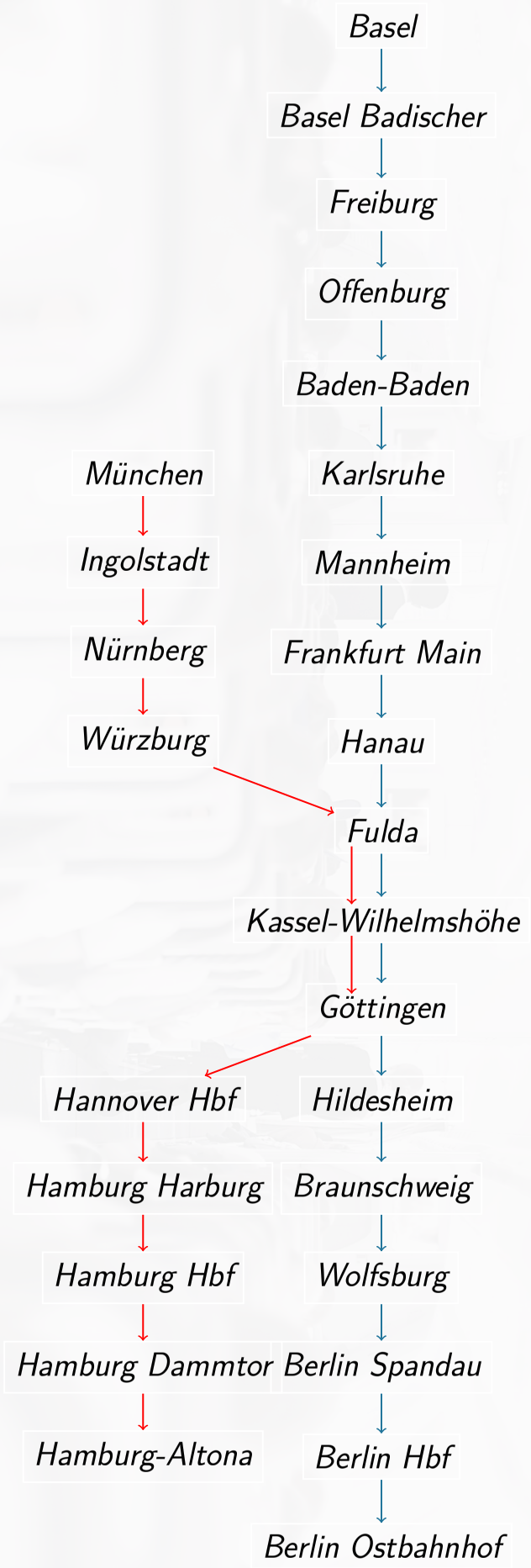
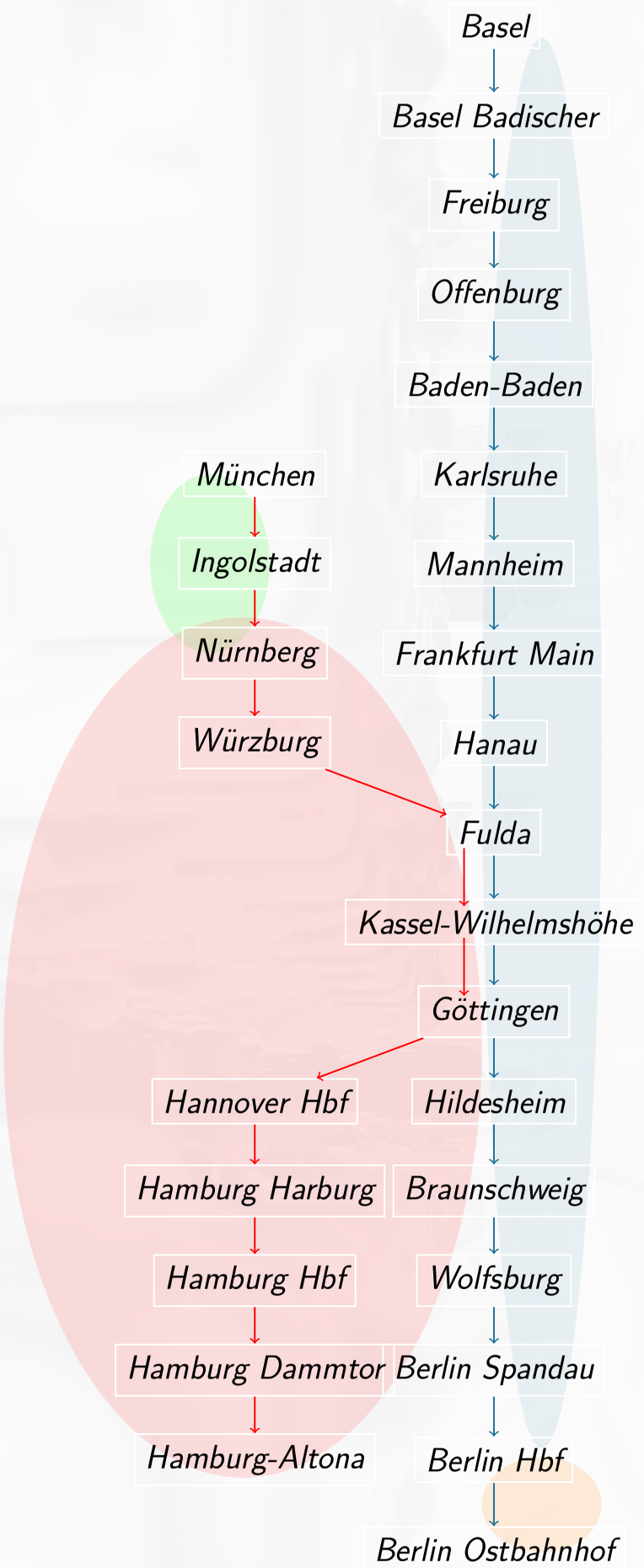
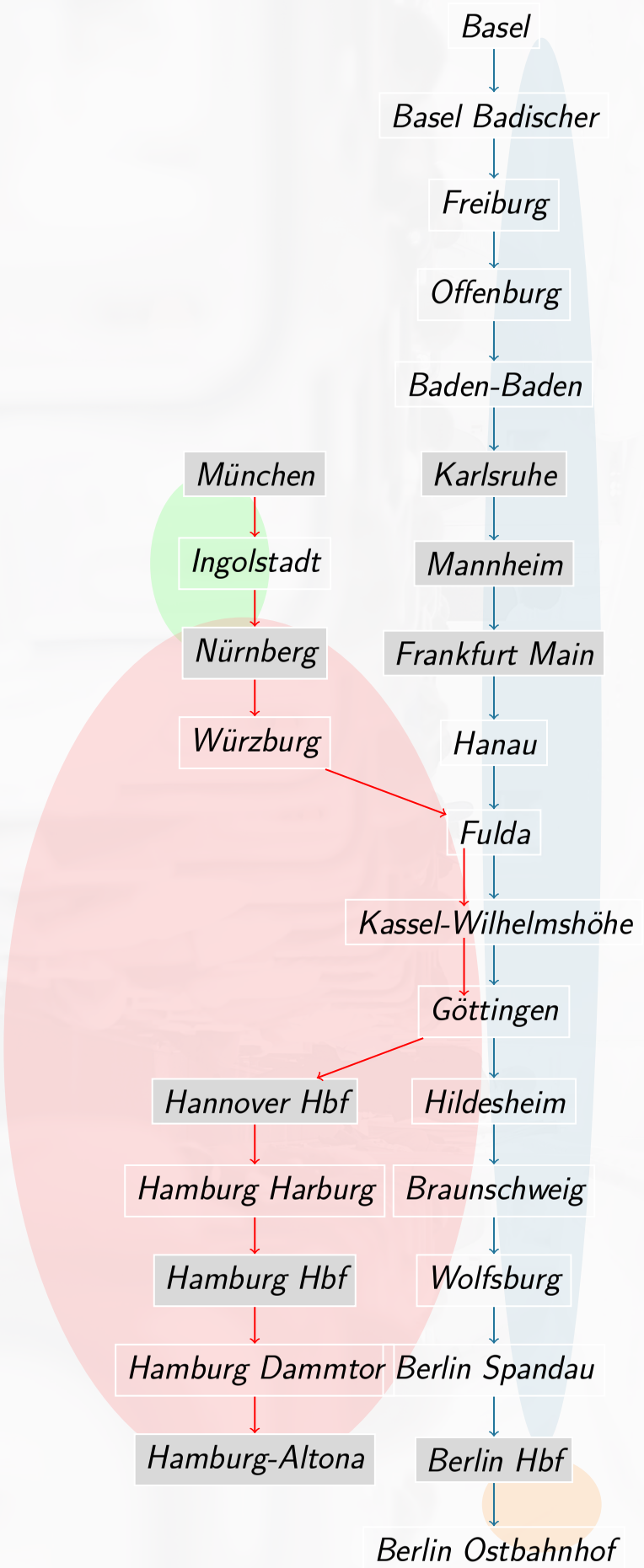
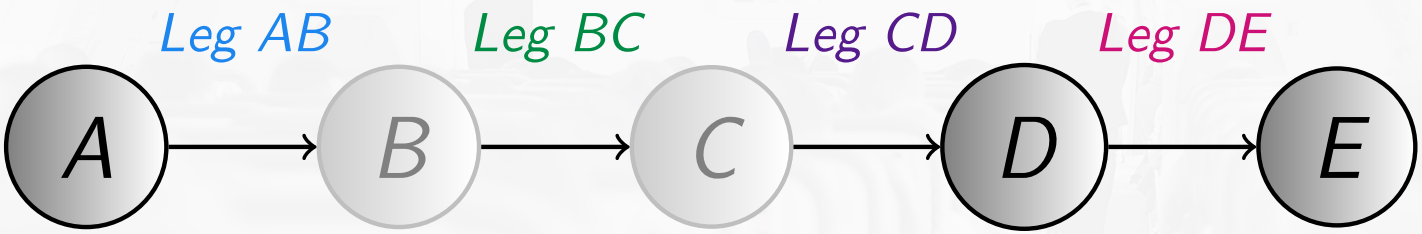


Figure: Clusters obtained in inverted graph


Basel
Basel Badischer
Freiburg
Offenburg
Baden-Baden
Karlsruhe
Mannheim
Frankfurt Main
Hanau
München
Ingolstadt
Nürnberg
Würzburg
Fulda
Kassel-Wilhelmshöhe
Göttingen
Hannover Hbf
Hildesheim
Hamburg Harburg
Braunschweig
Hamburg Hbf
Wolfsburg
Hamburg Dammtor
Berlin Spandau
Hamburg-Altona
Berlin Hbf
Berlin Ostbahnhof


Basel
Basel Badischer
Freiburg
Offenburg
Baden-Baden
Karlsruhe
Mannheim
Frankfurt Main
Hanau
Fulda
Kassel-Wilhelmshöhe
Göttingen
Hildesheim
Braunschweig
Wolfsburg
Berlin Spandau
Berlin Hbf
Berlin Ostbahnhof
München
Ingolstadt
Nürnberg
Würzburg
Hannover Hbf
Hamburg Harburg
Hamburg Hbf
Hamburg Dammtor
Hamburg-Altona


Basel
Basel Badischer
Freiburg
Offenburg
Baden-Baden
Karlsruhe
Mannheim
Frankfurt Main
Hanau
Fulda
Kassel-Wilhelmshöhe
Göttingen
Hildesheim
Braunschweig
Wolfsburg
Berlin Spandau
Berlin Hbf
Berlin Ostbahnhof
München
Ingolstadt
Nürnberg
Würzburg
Hannover Hbf
Hamburg Harburg
Hamburg Hbf
Hamburg Dammtor
Hamburg-Altona


Leg AB
Leg BC
Leg CD
Leg DE
A
B
C
D
E

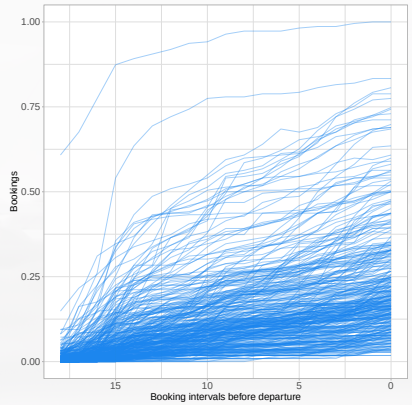


Figure: Leg AB

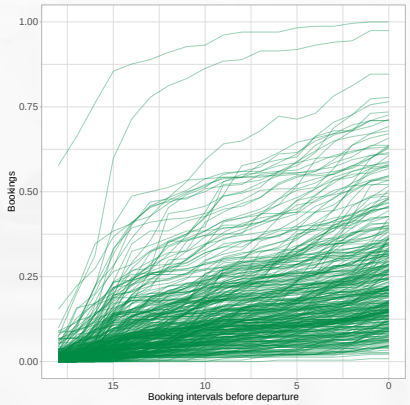


Figure: Leg BC

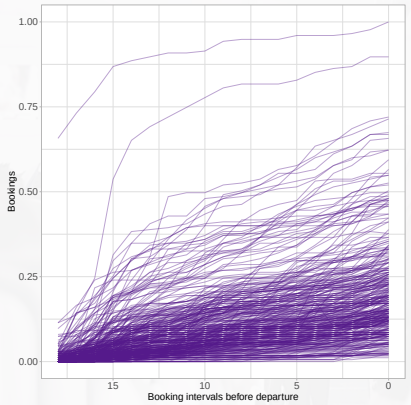


Figure: Leg CD

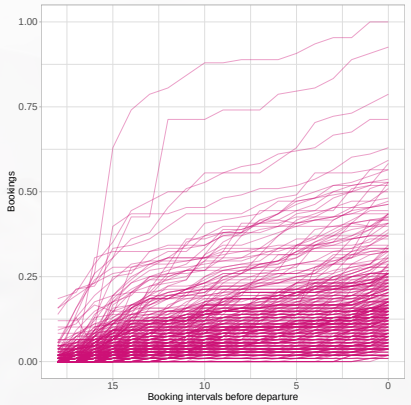
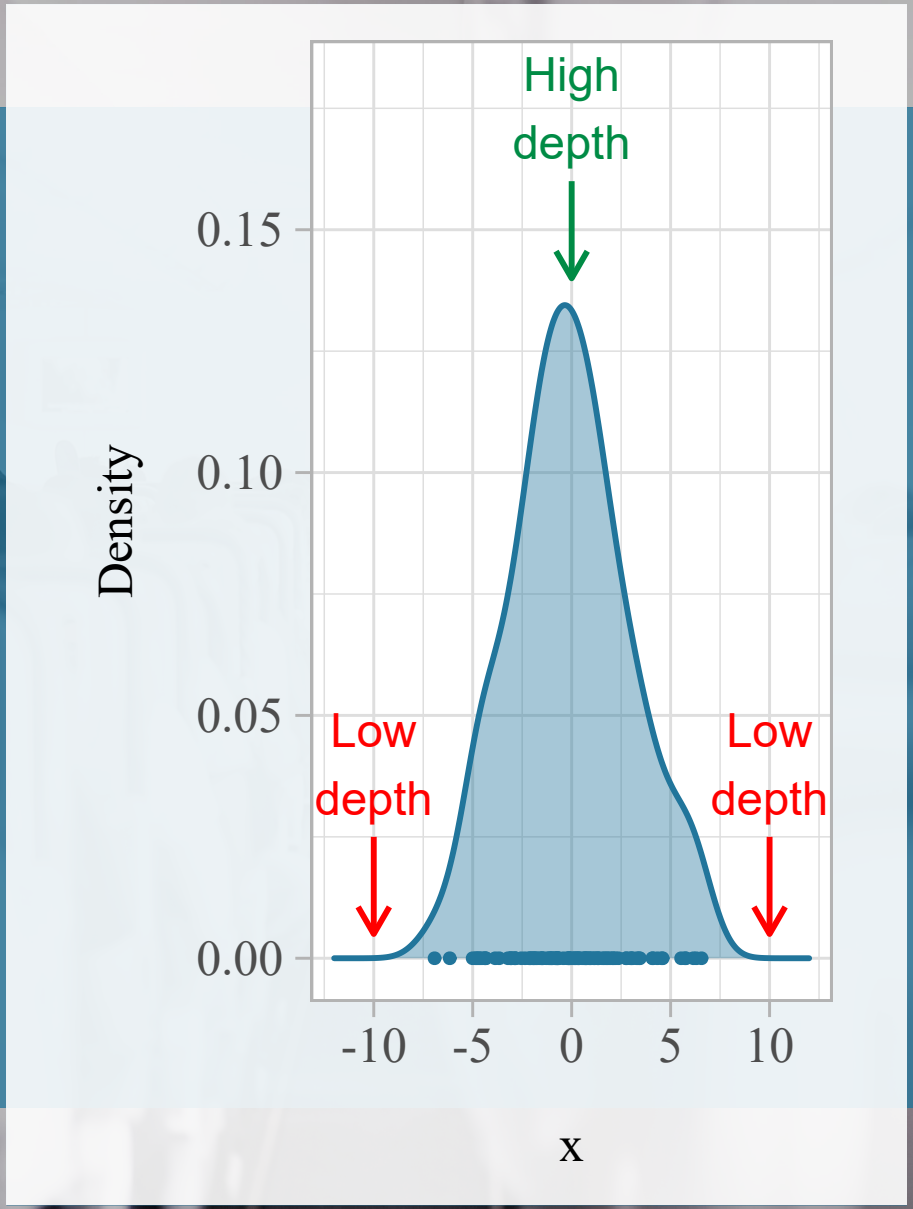
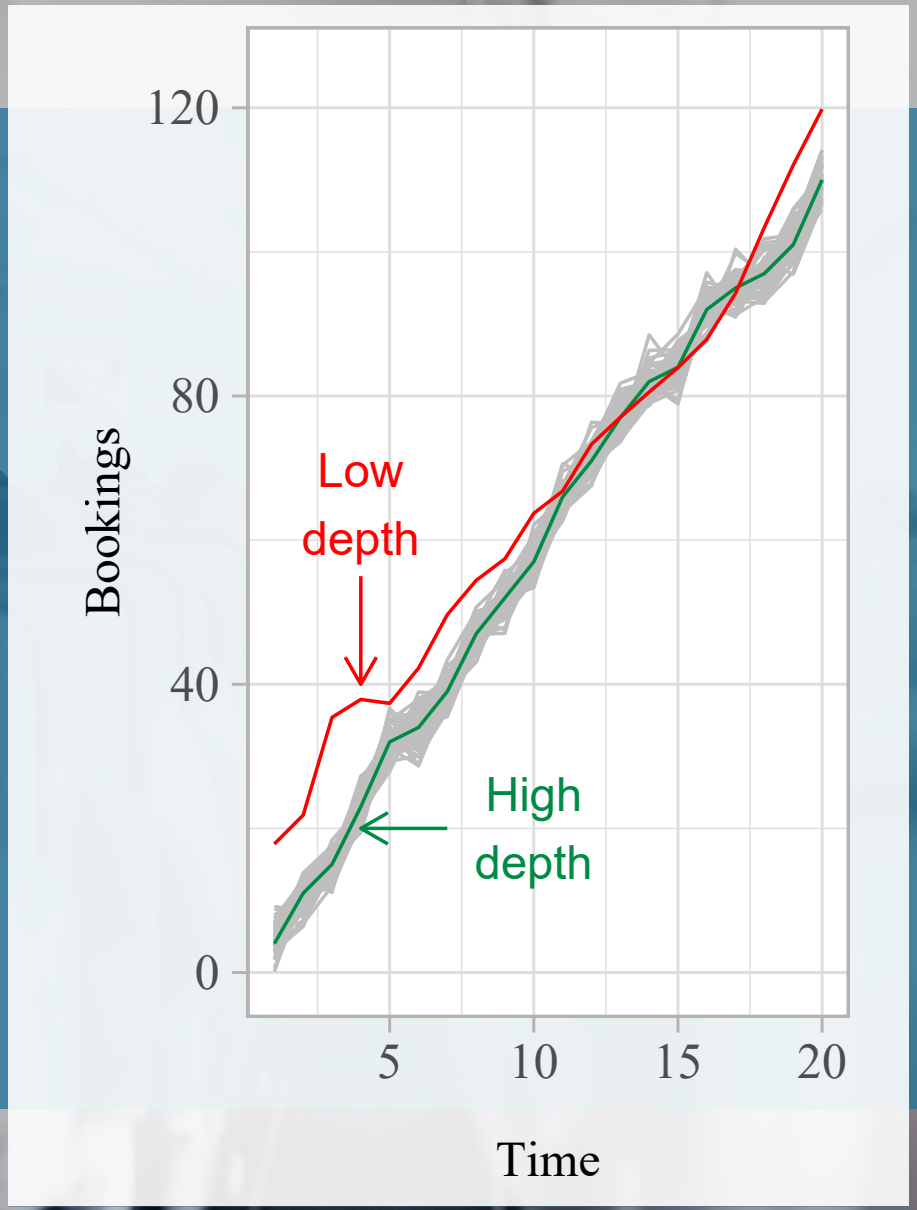


Figure: Leg DE

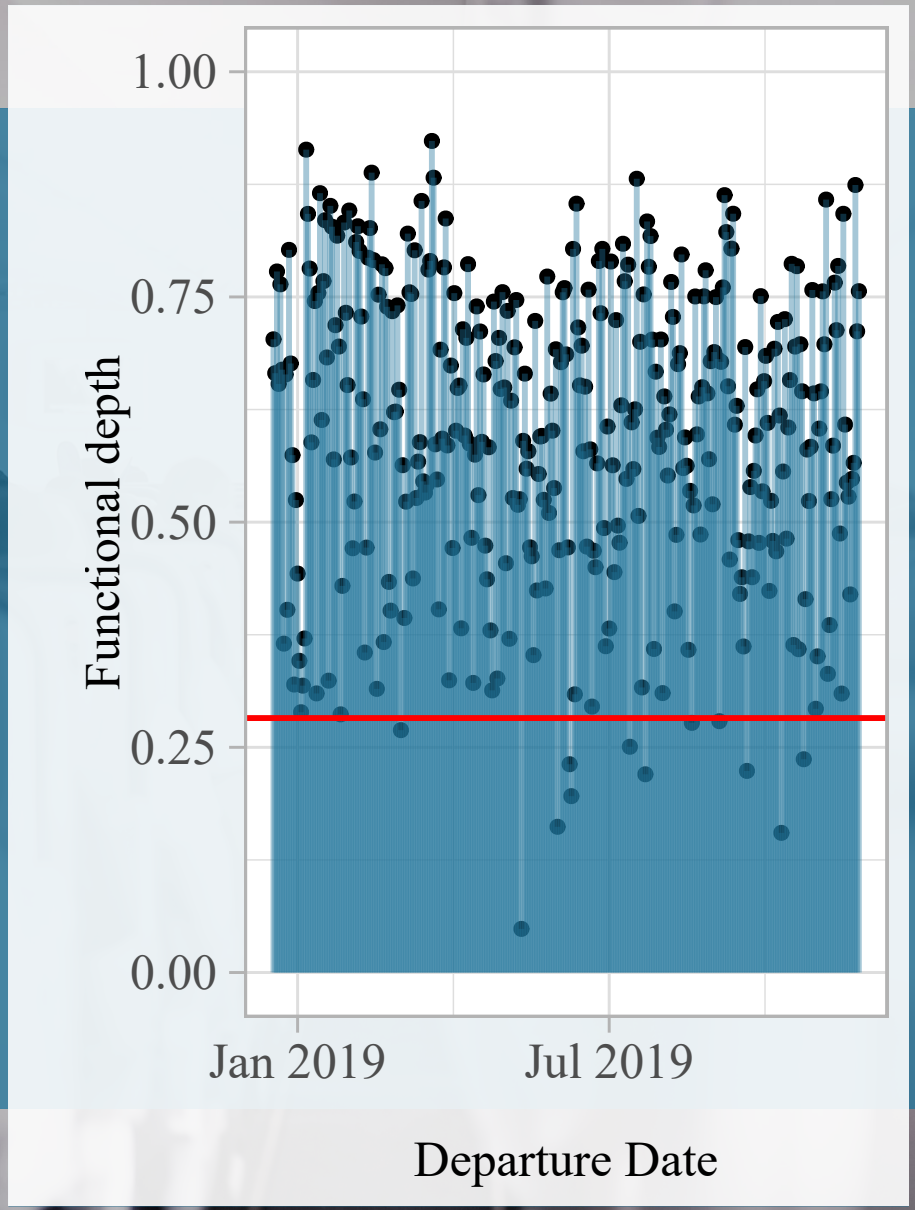
- *Depth measures provide an ordering of the data*



- ***Functional** depth quantifies how central a trajectory is*

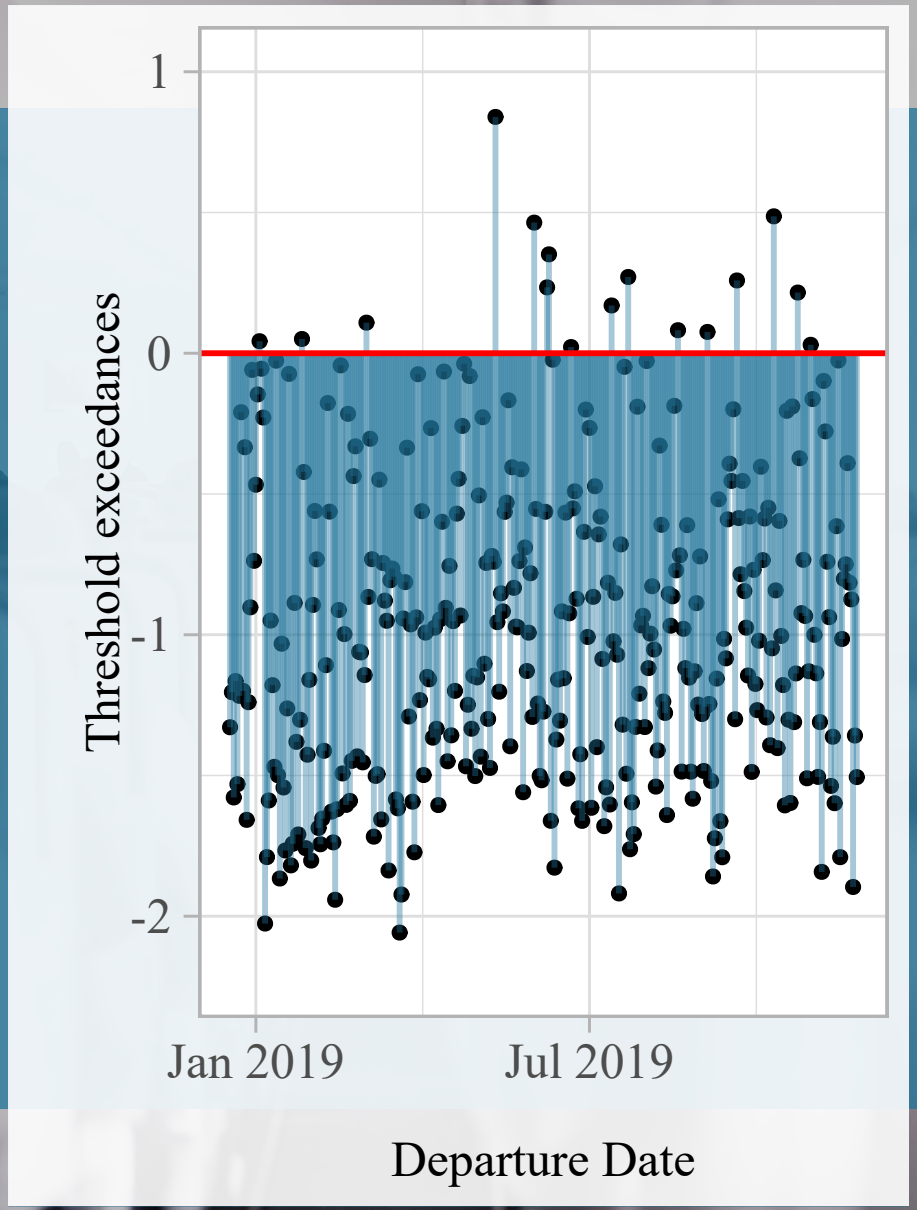


- *Define a threshold for the functional depth on each leg*
- *Departures with depth below threshold are outliers*



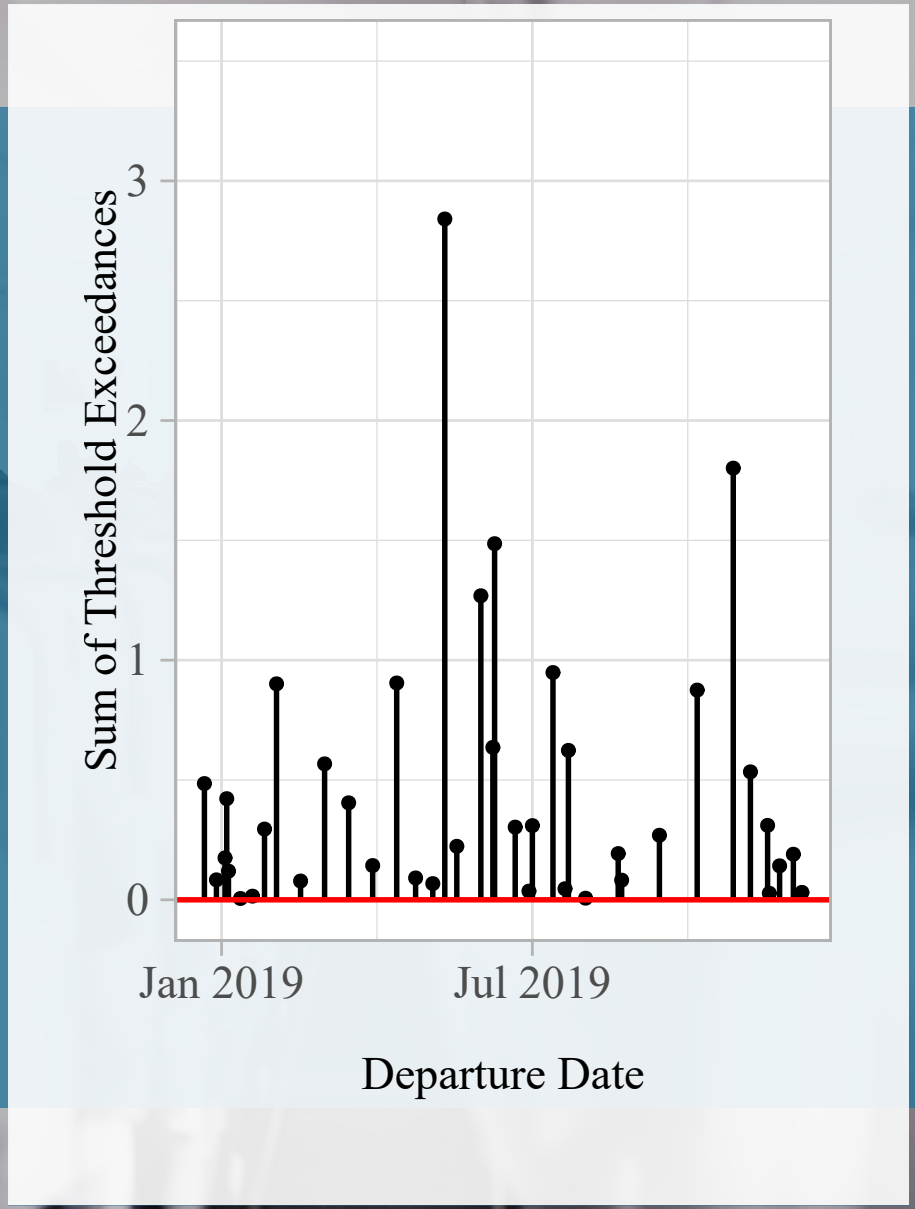
- *Define z_{nl} to be the normalised difference between the functional depth and the threshold:*

$$z_{nl} = \frac{C_l - d_{nl}}{C_l}.$$

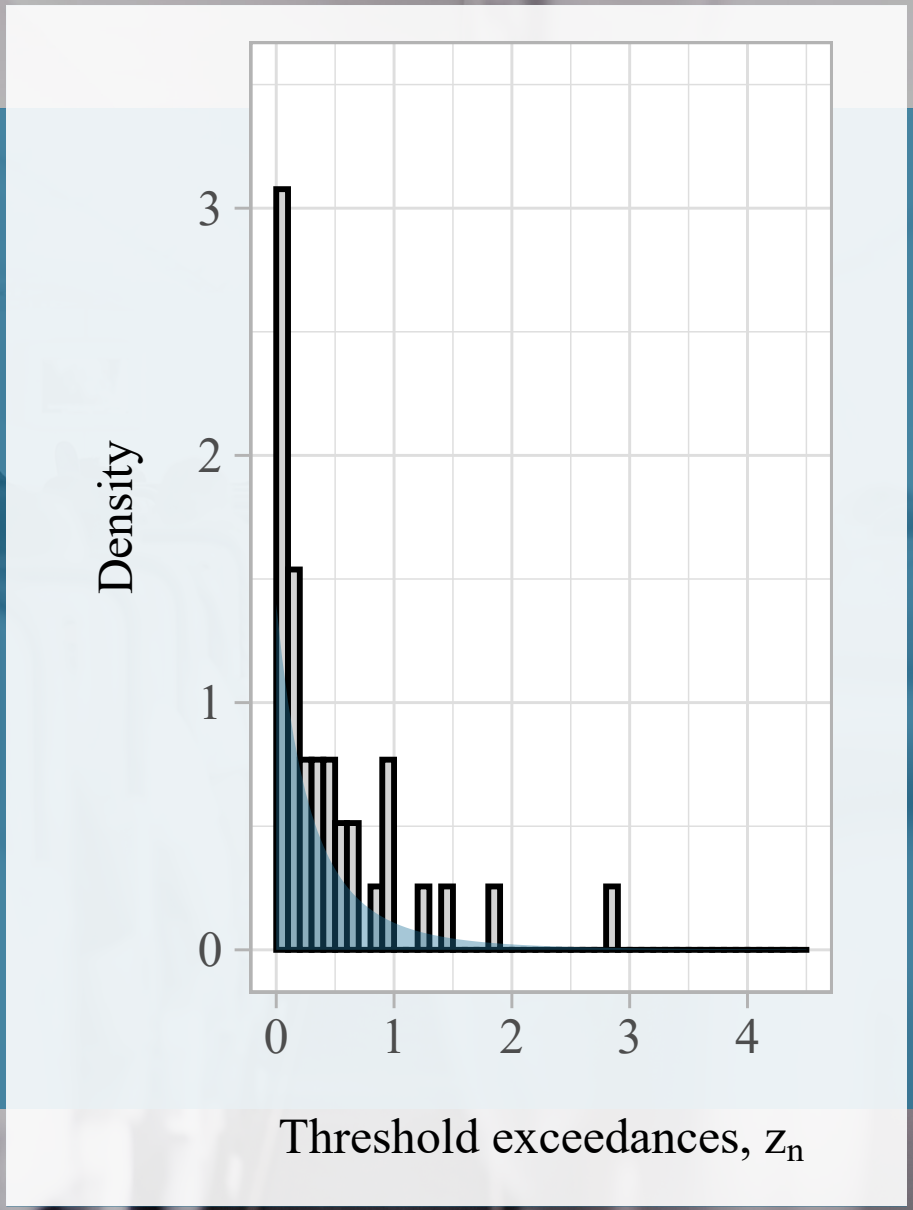


- *Next we define the sums of threshold exceedances across legs:*

$$z_n = \sum_{l=1}^{L} z_{nl} \mathbb{1}_{\{z_{nl}>0\}}.$$



- *We want to measure outlier severity.*
- *Fit a distribution to the threshold exceedances.*



Define θ_n to be the non-exceedance probability from the GPD. The non-exceedance probability is given by the CDF:

$$\theta_n = F_{(\mu,\sigma,\xi)}(z_n) = \begin{cases} 1 - \left(1 + \frac{\xi(z_n - \mu)}{\sigma}\right)^{-\frac{1}{\xi}} & \xi \neq 0 \\ 1 - exp\left(-\frac{(z_n - \mu)}{\sigma}\right) & \xi = 0 \end{cases}$$

Construct an alert list to send to analysts:

Ranking	***Departure***	***Probability***	***Legs Detected In***
1	*11/05/2019*	*0.985*	*AB, BC, CD, DE*
2	*26/10/2019*	*0.960*	*AB, BC, CD, DE*
3	*09/06/2019*	*0.942*	*AB, BC, CD, DE*
4	*01/06/2019*	*0.874*	*AB, BC, CD, DE*
5	*13/07/2019*	*0.865*	*AB, BC, CD, DE*
⋮	⋮	⋮	⋮

Table: Ranked alert list for cluster $= \{AB, BC, CD, DE\}$

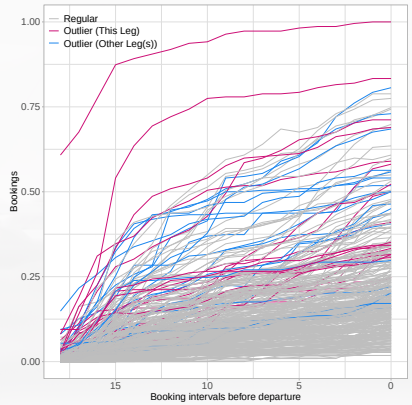


Figure: Leg AB

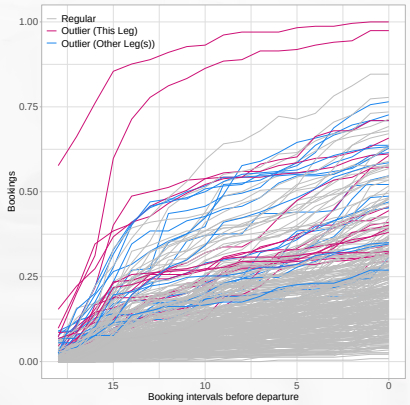


Figure: Leg BC

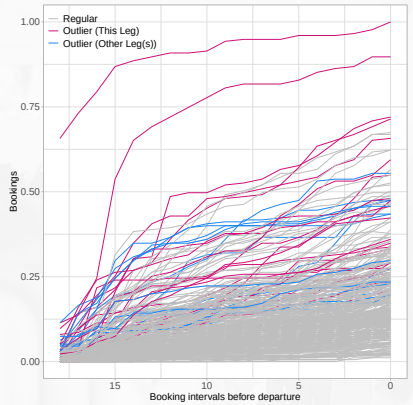


Figure: Leg CD

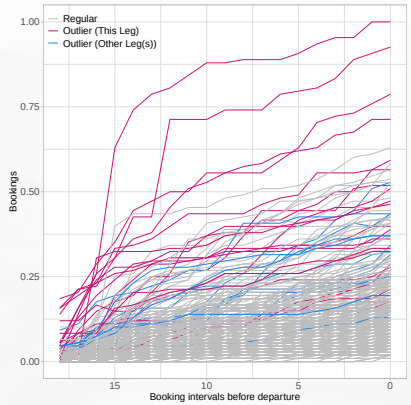
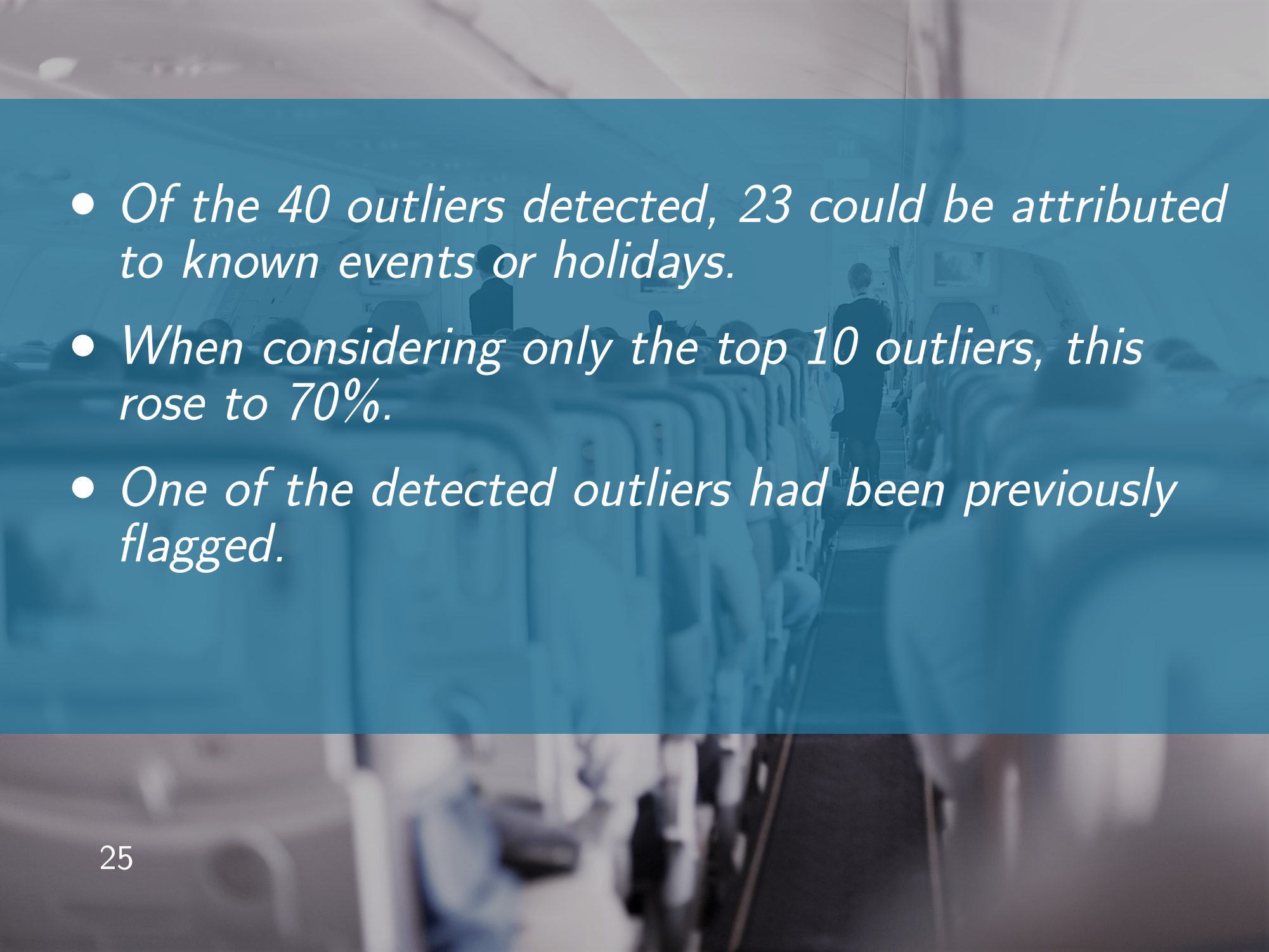


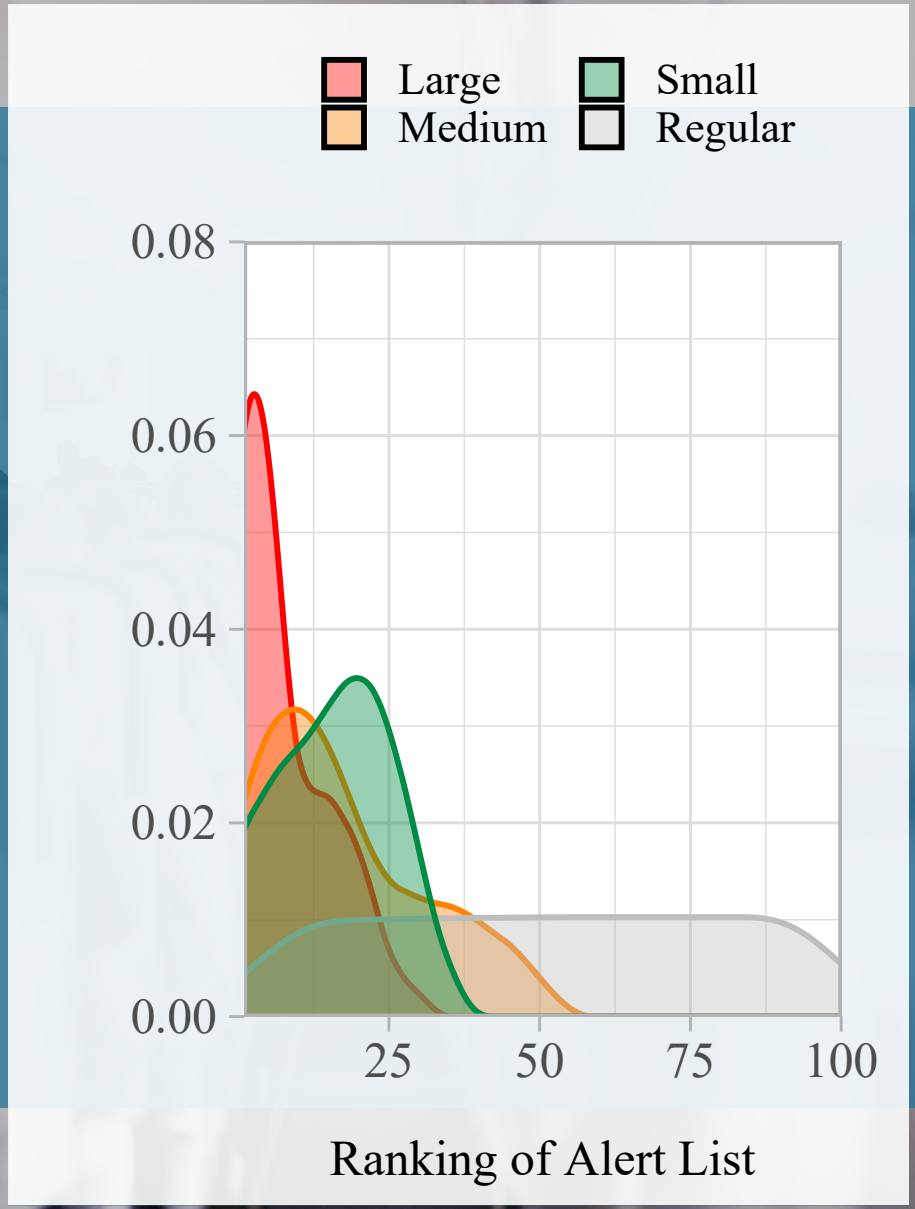
Figure: Leg DE



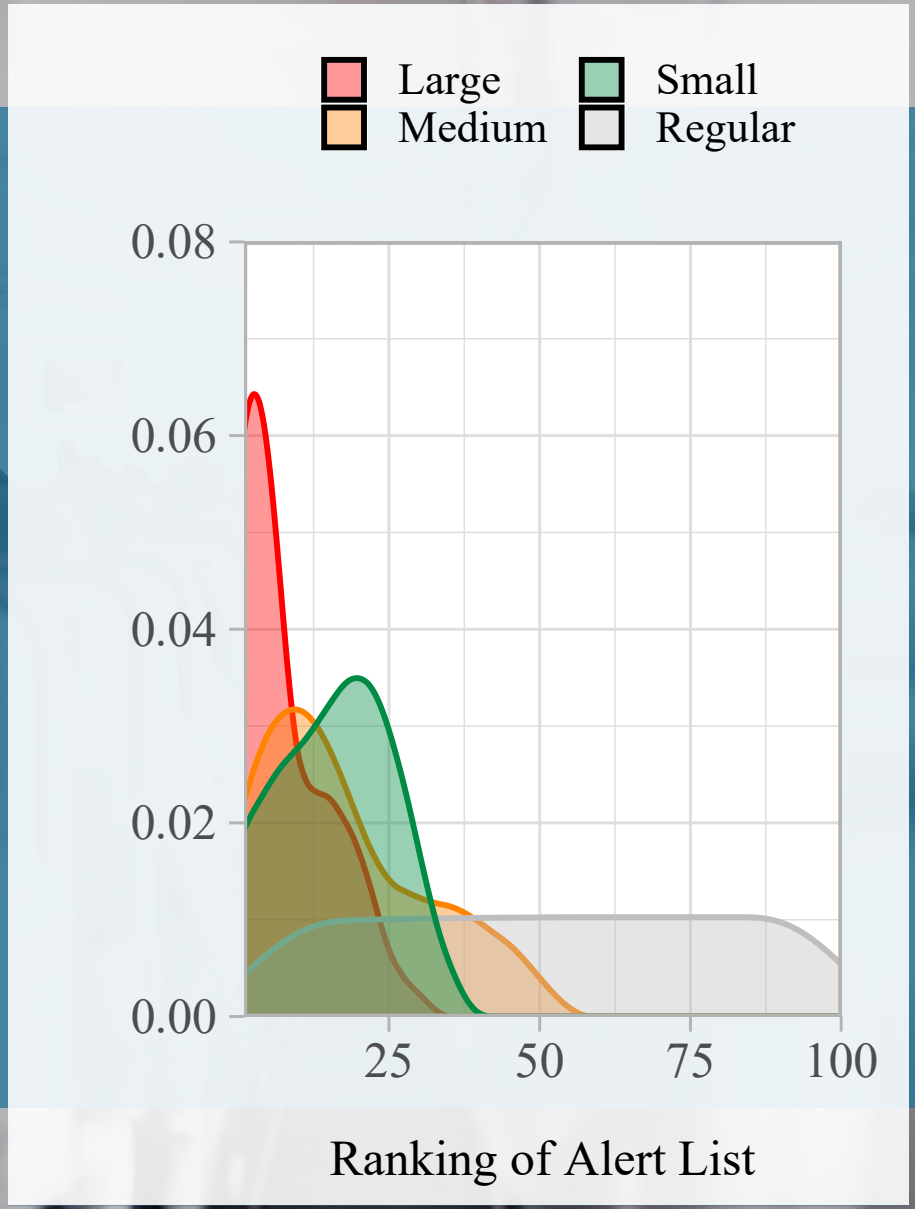
- *Of the 40 outliers detected, 23 could be attributed to known events or holidays.*
- *When considering only the top 10 outliers, this rose to 70%.*
- *One of the detected outliers had been previously flagged.*

Does it work?

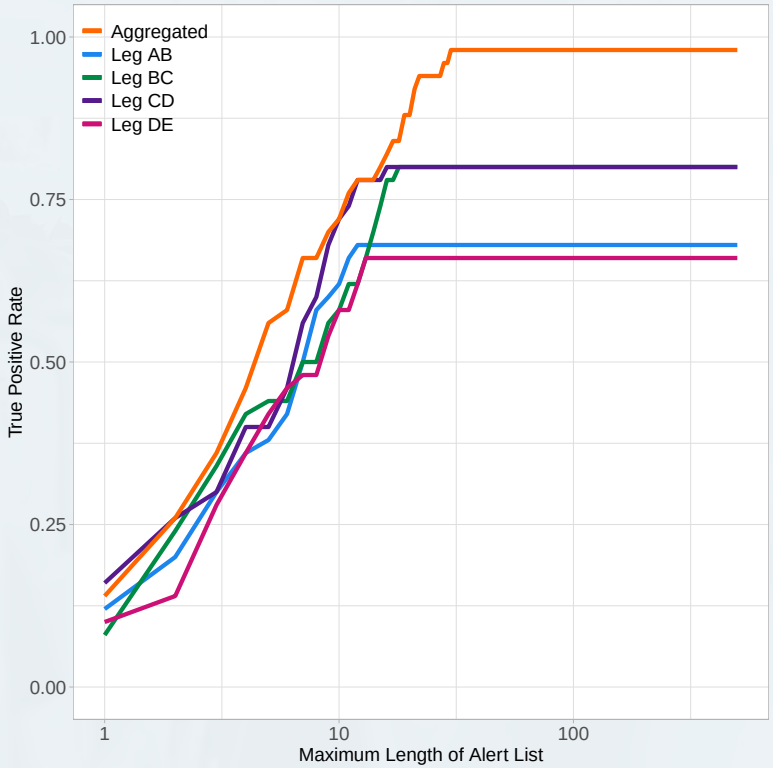
- *Use simulation to evaluate detection and ranking of outliers*



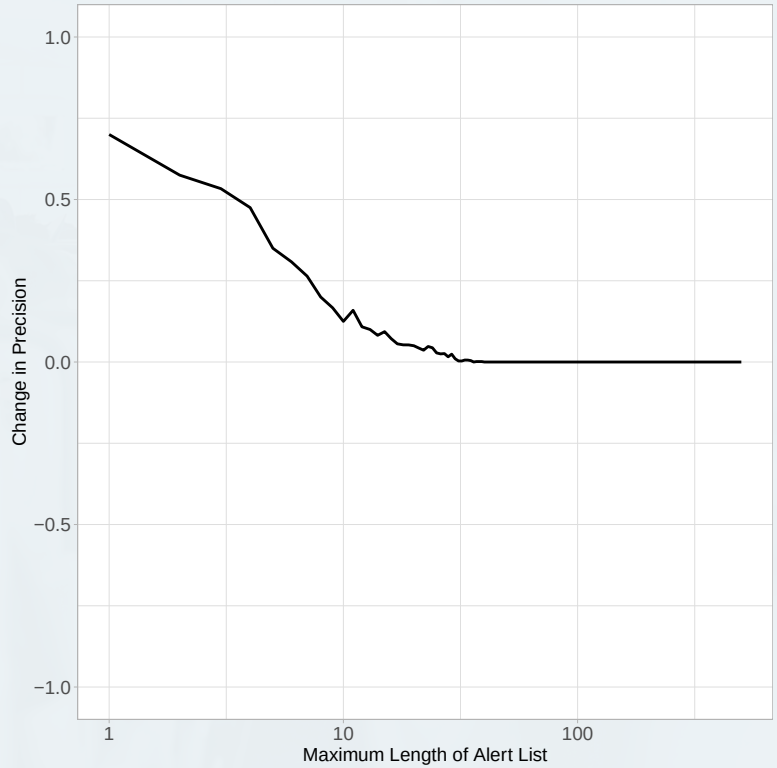
- *Larger outliers are ranked higher*
- *Ranking of medium outliers depends on sizes of other outliers*

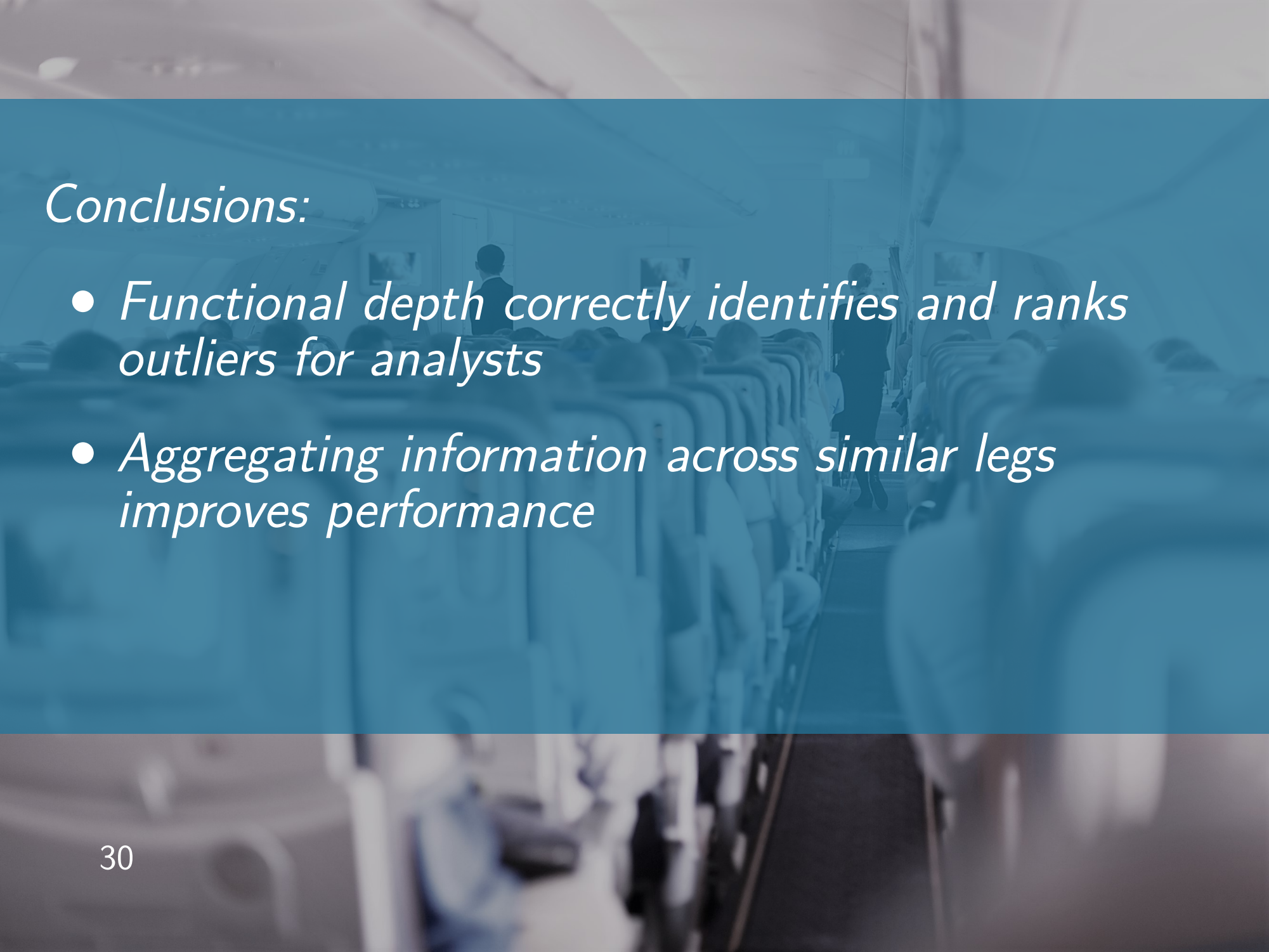


- ***True positive rate**: fraction of genuine outliers which have been detected.*



- ***Precision***: *fraction of departures classified as outliers that are genuine outliers.*





Conclusions:

- *Functional depth correctly identifies and ranks outliers for analysts*
- *Aggregating information across similar legs improves performance*

- *N. Rennie, C. Cleophas, A.M. Sykulski et al. Identifying and responding to outlier demand in revenue management. European Journal of Operational Research. 2021.*
- *N. Rennie, C. Cleophas, A.M. Sykulski et al. Detecting outlying demand in multi-leg bookings for transportation networks. arXiv. 2021.*

n.rennie@lancaster.ac.uk • @nrennie35