



DataMOCCA

Data MOdel for Call Center Aalysis

Project Collaborators:

Technion: P. Feigin, A. Mandelbaum

SEELab: V. Trofimov, E. Nadjharov, I. Gavako, K. Kutsy;

P. Khudyakov, P. Liberman, A. Senderovic, S. Maman;

Students (PhD, MSc, BSc), RAs

Wharton: L. Brown, N. Gans, H. Shen (N. Carolina),
Students, Wharton Financial Institutions Center

Industry: U.S. Bank, Israeli Telecom, 2 Israeli Banks,
Israeli Hospitals, ..., IBM Research

Goal: Support Data-Based Research & Teaching

i.e. Designing and Implementing a (source-independent) **transaction**-based data repository, which is (universally) accessible within a graphical **online** EDA environment: cleaning, storing, retrieving, displaying, interacting, analyzing, **learning**, enjoying.

The Data Challenge - An Example

- Mid 90's: Industry lectures on significance of **transaction-based data**
- "Everyone could" but only one would - with **Technion doing the work!**
- Executed by a single undergraduate student + Technion Statistics Lab
- 1999, **Bank Anonymous**: Disappointingly (but Fortunately) only 15 Agents, 350K call/year

The Data Challenge - An Example

- Mid 90's: Industry lectures on significance of **transaction-based data**
- "Everyone could" but only one would - with **Technion doing the work!**
- Executed by a single undergraduate student + Technion Statistics Lab
- 1999, **Bank Anonymous**: Disappointingly (but Fortunately) only 15 Agents, 350K call/year
- Wharton team (**Brown, Gans**) joined and helped recruit U.S. Bank: but now 350K call per week, over 700 agents
- Attempts by a Wharton PhD student + some help, but too much
- Cutting a long story short: with **Paul Feigin**, created **SEELab** in 2007, through the generous gift of **Hal and Inge Marcus** (\$1M to IE at both Penn State and Technion) - recruited by Technion IE dean **Boaz Golany**

The Data Challenge - An Example, Continued

- 2008, **10 years later**: same industry lecture attended by Bank Anonymous (previously agents now) managers: data nostalgia helped rekindle the romance
- Another long story made sort: March **2008** - first meeting, January **2010** - first data arriving to the SEELab
- Rewards of patience: **daily data deposit** at the SEELab safe; plans to expand from operational- to also **business-data**

1. **Clean Databases:** Operational histories of **individual** customers and servers (mostly with IDs)
 - In Call Centers: from IVR to Exit
 - In Hospitals: from ED to Exit (or just ED)
2. **SEESat: Online EDA** (friendly, flexible, powerful)
 - Queueing-Science perspective
 - Operational data (vs. financial, contents or clinical)
 - Flexible customization (e.g. seconds to months)
3. **Tools:**
 - Online statistics (survival analysis, mixtures, smoothing);
 - Dynamic Graphs (flow-charts, work-flows) - in process
 - Simulators (CC, ED; data-driven) - in process