

#DF15DE



Herzlich willkommen zum
DataFest Germany!



ASA DataFest™ @Duke

[ABOUT](#)

[SIGN UP](#)

[SPONSORS](#)

[WORKSHOPS](#)

[SCHEDULE](#)

[MEDIA](#)

[FAQ](#)

[CONTACT](#)



ASA DATAFEST™ @DUKE

Are you data driven?

#datafestATduke

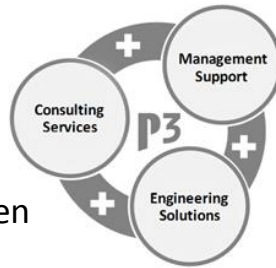
Duke / UNC / NCSU

March 20 - 22, 2015

Ipsos und P3 sammeln gemeinsam Daten in einem mobilen und agilen Umfeld

P3 GROUP

- P3 group bietet technische Beratung, Management Unterstützung, Entwicklungsleistungen und Test Dienstleistungen
- Branchen sind Telekommunikation, Automotive und Aviation
- P3 stößt in neue Märkte vor wie Energie und Bahn



P3 COMMUNICATIONS GMBH

- Beratung und Engineering im Mobilfunkbereich
- Bewertung und Optimierung von Mobilfunknetzen
- Connect Test
- Crowd sourced App basierte Bewertung der Mobilfunkqualität

IPSOS

- Ipsos liefert Unternehmen weltweit datenbasierte Entscheidungshilfen für Geschäftsentscheidungen
- Antworten auf Kundenfragen liefern fünf Spezialisierungen: Ipsos Marketing, Ipsos Loyalty, Ipsos Public Affairs, Ipsos Connect, Ipsos Operations
- Past: Konsumentenbefragungen

TECHNOLOGIEBASIERTE VERHALTENSMESSUNGEN

- Present – Future: Digitalisierung, Mobilisierung ermöglichen den verstärkten Einsatz von Technologien zur Verhaltensmessung und liefern damit ein genaueres Bild, was Konsumenten tun
- Durch die P3 Applikation verstehen wir, was Smartphone-Nutzer auf ihrem Gerät machen



Passive erfasste Daten von Smartphones während der täglichen Nutzung

IPSOS PANEL

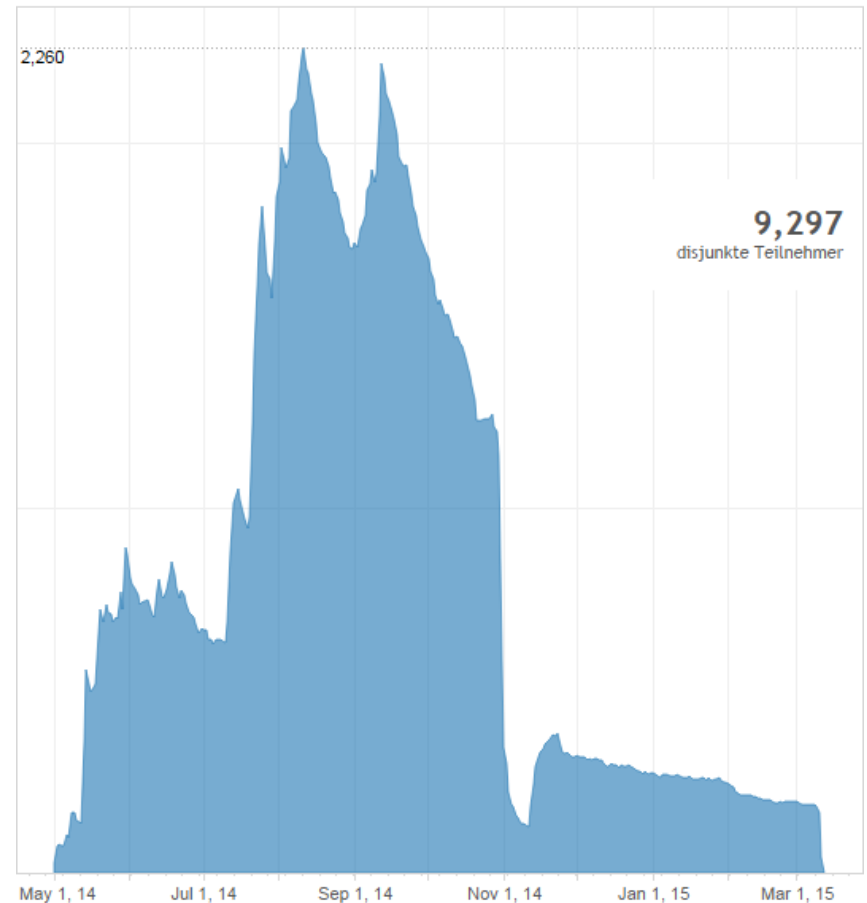
- Ipsos betreibt Panels mit freiwilligen Teilnehmern und führt passive Untersuchungen durch
- Die Teilnehmer installieren eine Tracking App auf ihren Smartphones
- Mobile passive Datenerfassung erlaubt Ipsos Daten in real-time in einem sehr agilen Umfeld zu erfassen
- Es werden Daten gesammelt wie, wo und wann die Teilnehmer ihr Smartphone nutzen

DIE GESAMMELTEN DATEN ENTHALTEN

- Besuchte Webseiten
- Installierte Apps und deren Nutzung
- Aufenthaltsort während der Nutzung
- Zeit und Datumstempel
- Demografie
- Device Typen/Hersteller
- ...



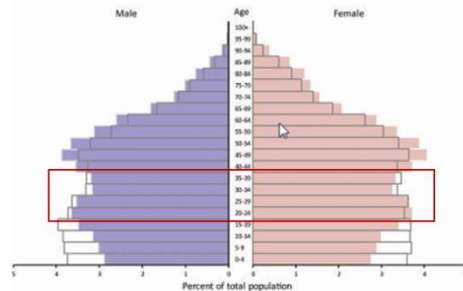
ANZAHL TEILNEHMER



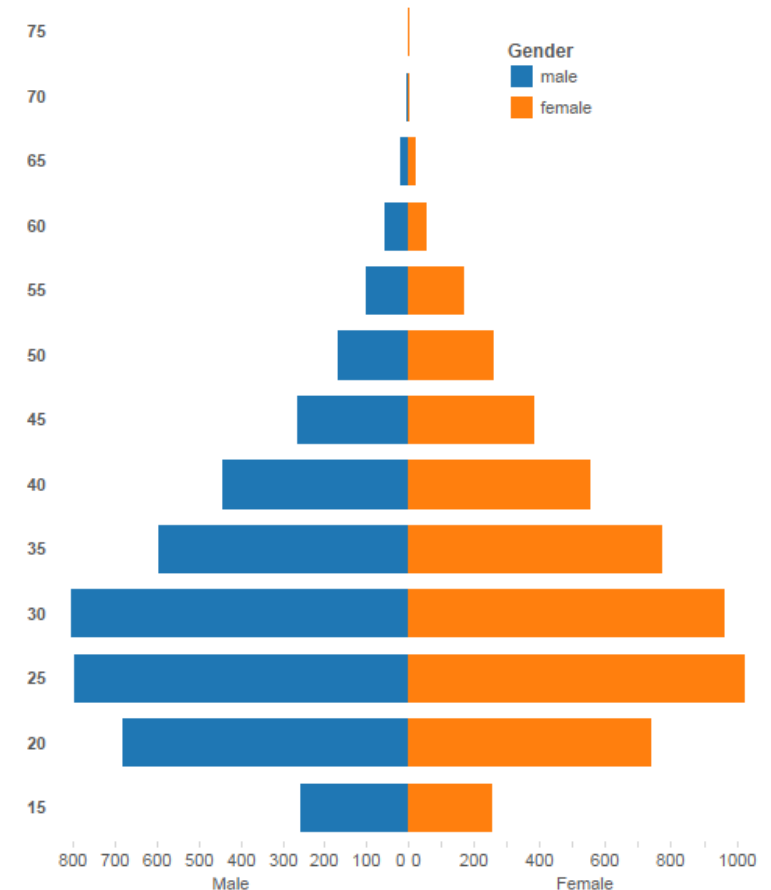
Bei der Rekrutierung werden demografische Daten erhoben

FELDHINHALTE

- Alter
- Geschlecht
- Beschäftigung
- ExternalID
als eindeutige ID
- Gerätetyp
Tablet vs Telefon
- Mobilfunk relevante Größen
 - Netzbetreiber
 - Inkludiertes Datenvolumen
 - Vertragstyp (Prepaid/Postpaid)



ALTERS PYRAMIDE



Jeder Start einer Smartphone App erzeugt ein Datum

FELDHINHALTE

- Start und Ende einer Nutzung erzeugt ein Datum
 - App Name
 - Package Name
 - AppStart/EndTime
jeweils UTC
 - UserHash als eindeutige ID

SAMPLE (HIER NUR 30 USERS UND 24H)



Dieser Datensatz liefert Zusatzinformationen zu den Teilnehmern und verknüpft UserHash und ExternalID

FELDHALTE

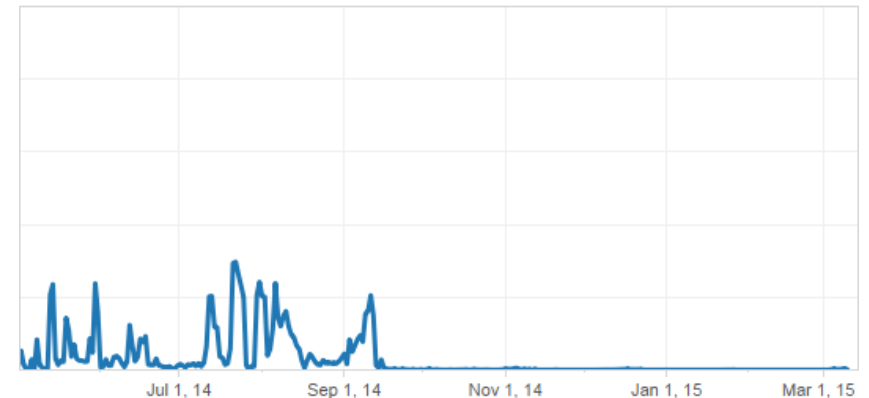
- App Registration Time
- App Version
- Device Manufacturer
- Device Model
- OS
- OS Version
- ExternalID
- UserHash

ANZAHL TEILNEHMER/BRUTTO ADD

Distinct Users



Brutto Add

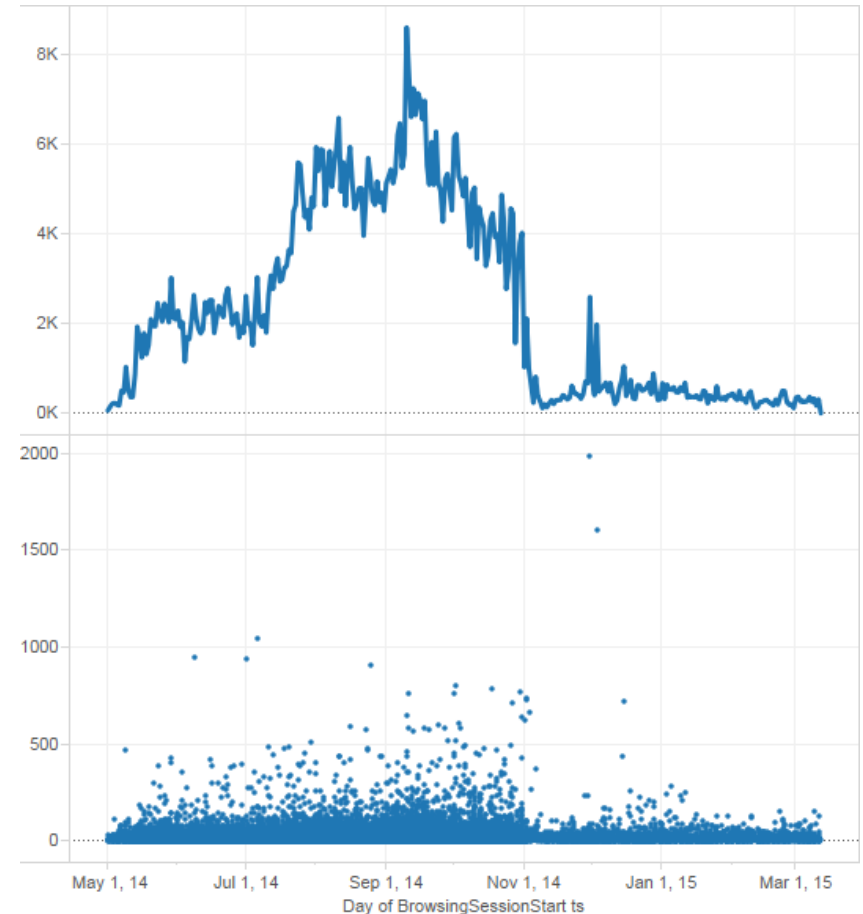


Die auf dem nativen Browser und in Google Chrome aufgerufenen URLs

FELDHINHALTE

- Bookmarked
- SessionStart/End
- Last Visited
- Title
- URL
- UserHash

ANZAHL URL PRO TAG BZW PRO USER

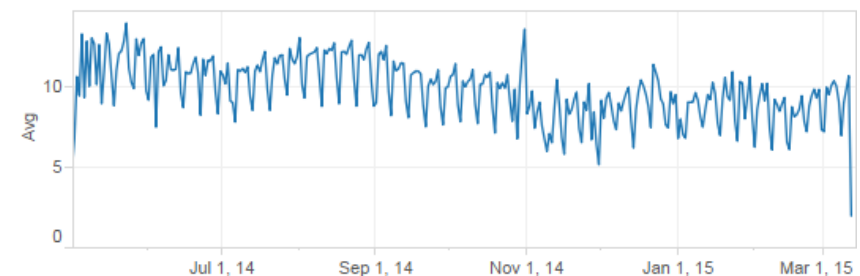
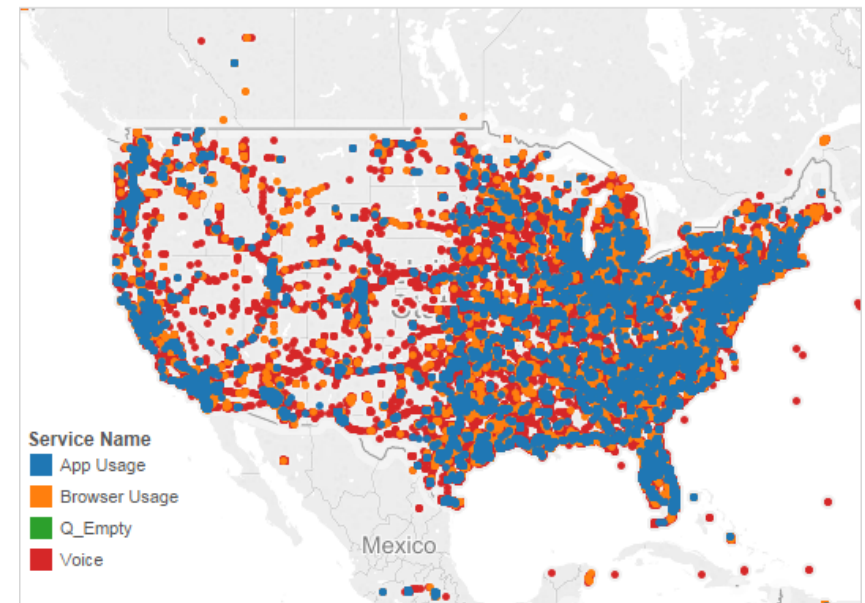


Die Nutzung bestimmter Apps und Telefongespräche lösen auch eine Positionierung aus

FELDHINHALTE

- Die Lokalisierung durch GPS ist nur mit erheblichem Strombedarf durchzuführen
- Deshalb wurde die Lokalisierung nicht bei jeder App Nutzung durchgeführt
- Lokalisierung wurde durch Telefonie, Browser Nutzung und bestimmte Apps getriggert (z. B. YouTube, Flipboard, eBay, Amazon, Shazam, Spotify, ...)
- Durchschnittlich >10 Orts-/Zeitstempel pro User und Tag

ÖRTLICHE VERTEILUNG



Real-world user profiles with demographic, geographic, and behavioral information generated by analyzing geo-behavioral patterns



INHALTE

- Factual ermittelt aus Geo-Traces (Lat/Long/Zeit) Demografische Informationen
- Factual bekam im Vorfeld dieser Veranstaltung das Location Dataset
- Daraus wurden die Zusatzinformationen der einzelnen Teilnehmer berechnet
- Eine detaillierte Beschreibung findet sich auf dem USB Stick, den Sie bekommen

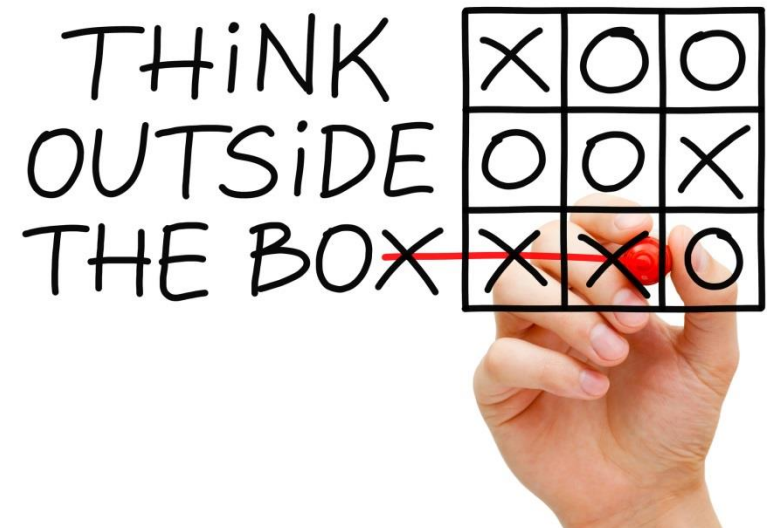
SAMPLE

```
{
  "age" {
    "range" "18-24"
    "confidence" 0.75
    "index" 2
  }
  "family" {}
  "race" {
    "african_american" {
      "confidence" 0.04
    }
    "hispanic" {
      "confidence" 0.03
    }
  }
  "behavioral" {
    "segments" {
      "qsr" {
        "confidence" 0.97
        "percentile" 98.5
      }
      "date_nighter" {
        "confidence" 0.79
        "percentile" 89.4
      }
      "shopping_enthusiast" {
        "confidence" 0.89
        "percentile" 94.5
      }
      "financial_customer" {
        "confidence" 0.88
        "percentile" 94.1
      }
      "affluent_consumer" {
        "confidence" 0.97
        "percentile" 98.5
      }
    }
  }
  "interests" {}
}
```

Aktuelle Fragestellungen, die im Bereich Mobiler App und Web Nutzung interessieren

FRAGEN IM KONTEXT MOBILE APP WEB NUTZUNG

- Der Mobile User: Gibt es Unterschiede nach Alter & Geschlecht? Unterschiede nach Jahreszeiten? Events?
- Der App User: Gibt es Unterschiede nach Alter & Geschlecht?
- Gibt es eine Verbindung zwischen App-Nutzung und dem Besuch eines Ortes?
- Wie lässt sich zuhause vs. unterwegs analysieren? Wie gestaltet sich die Nutzung zuhause vs. unterwegs? Anteile – Frequenz – Inhalte?
- Für einen spezielle Kunden: Wie sind die Anteile von Mail-Programmen wie gmail, Hotmail, yahoo mail und anderen Apps und Webmail Angeboten?



Datenschutz ist ein Thema mit dem vertrauensvoll umgegangen werden muss

WARUM IST ES WICHTIG FÜR UNS

- P3/Ipsos haben mit den Nutzern Terms&Conditions unter denen die Teilnehmer ihre Daten bereitstellen
- Wir haben zusätzliche Verschwiegenheitsvereinbarungen unterzeichnet um euch den Zugriff auf diese Daten zu ermöglichen
- Die Daten dürfen außerhalb dieser Veranstaltung nicht verwendet werden
- Alle Kopien müssen nach Ende der Veranstaltung gelöscht werden
- Erstellte Präsentationen/Analysen dürfen unter der Angabe von Quellen verwendet werden

WAS TUN WENN ES DOCH PASSIERT

- Ruhe bewahren
- Nichts vertuschen
- Veranstalter informieren



WARUM IST ES WICHTIG FÜR EUCH

- Weil es für uns wichtig ist
- Weil ihr ein NDA unterschrieben habt
- Ihr werdet im Data Science Bereich für Firmen arbeiten die euch Zugriff vertraulichen Daten geben
- Mit diesem Vertrauen müsst ihr umgehen lernen
- Firmen die im Bereich Data Science einstellen haben auch überdurchschnittliche Fähigkeiten im Bereich der Datenanalyse
- Die suchen nicht nur auf Facebook nach Bildern
- Sollte euer Name in Verbindung mit einem Data Leak stehen ist das sicher schädlicher als ein „NSFW“ or „TP“ Bild



Source: <https://www.flickr.com/photos/37606287@N00/2647269314/>

NSFW: Not suitable for work; TP: too private

Ablauf

Freitag

18:00	Check-In
19:00	Meet & greet the data – Raum A 001
20:00	Abendessen
21:30	TechTalk: Einführung in R und Json
00:00	Midnight Surprise – Lounge 2.OG

Samstag

9:00	Frühstück
10:00	Datensprechstunde - Raum A 001
11:00	TechTalk: Einführung in Google Maps Daten
13:00	Mittagessen
14:00-16:00	Career Fair
18:30	Abendessen
22:00-24:00	SAP Midnight Challenge

Sonntag

9:00	Frühstück
11:45	Abgabe der Präsentationen – Raum A 001
12:00 – 13:30	Finale (4Min/Team) – Room A 001
13:30 – 14:00	Awards – Room A 001

Sponsoren und Mitwirkende

Sponsoren + Jury



UNIVERSITÄT
MANNHEIM



Teams

1. SozÖkonomEtrics
2. Unsupervised Miners
3. Top of the Bell curve
4. R-squared Mannheim
5. YastiR
6. Paranormal Distribution
7. R - Force 1
8. para-normal distribution
9. Data Cleaner
10. Sander-Storm
11. Datenprimaten
12. ROC'n Roll
13. Team Psycho
14. CHIntas
15. R don't know
16. Hertie Data Nerds
17. Lokomotive Kolmogorov
18. Data-Diggers