



**5 YEARS OF  
@CHARLOTTEHACKS**

**WE ARE BACK**

WHO MADE IT POSSIBLE



**WHAT BROUGHT US ALL HERE TONIGHT**

**COMMUNITY | COMMUNITY | COMMUNITY**

**logistics**

## getting dialed in

- SPECTRUM GUEST WIFI: DISCOVERYPLACEGUEST
- PASSWORD: no password
- **SERVER AND DATA ACCESS WILL BE GIVEN BY TECH COMMAND CENTER. YOU WILL BE GIVEN ONE ID PER TEAM.**

# #HACKATHONCLT



@hackathonclt



/hackathonclt



@charlottehacks



# agenda

Friday March 24, 2017

5pm – Registration

7:30pm – Opening remarks & business problem introduction

8:30pm – Release the hackers | Data & Tech Help Desks open in the Command Center

11pm – Liquid Nitrogen Slushies

12am – Midnight snack

# agenda

Saturday March 25, 2017

2am – Oculus VR+ tips & tricks

6am – Breakfast & morning announcements

7am – Shortlisting begins (Windows on Tryon room) | Checkout begins

9am – Museum opens | Booths open | Shortlisting round ends

10am – Mimemantics performance in the amphitheater

11am – Checkout stops

12pm – Lunch in amphitheater | Final Presentations begin (10 teams) | Sponsor booth breakdown

2pm – Awards ceremony

## if you haven't noticed

- **NEW LOCATION** – In case you've missed it, for those of you coming from UNCC, you're not on campus anymore...
- **THOSE BOOTHS THO** – be sure to check out the unconventional, maker-style sponsor booths tonight & tomorrow morning between 9 – noon.
- **ENTERTAINMENT/ MIMEMATICS** –
  - giant jenga
  - yard games
  - mimematics
  - hack & hunt
  - wine + nitrogen

# THINGS TO REMEMBER & DO

- **VOLUNTEERS**
- **PARTICIPANTS**
- **EGRESS**
- **HERE ALL NIGHT** – to compete, we are here. The. Whole. Time. (but don't worry, we're here right with you)
- **CHECK OUT**
- **TEAM OR NO TEAM** - reminder that these are teams of up to 3 and no more. AND IF YOU ARE EXCEPTIONALLY GOOD OR EXCEPTIONALLY GREEDY A TEAM OF ONE WORKS

# rules of engagement

- nothing illegal
- respect copyright
- keep it clean
- terms & conditions
- all work must be on site
- organizers reserve the right...

**the problem**

**introducing this year's event beneficiary – Big Brothers Big Sisters**



**Big Brothers Big Sisters**

## IMPACT OF MENTORING

**FOR MORE THAN 100 YEARS NATIONALLY AND 45 YEARS LOCALLY, BIG BROTHERS BIG SISTERS HAS OPERATED UNDER THE BELIEF THAT INHERENT IN EVERY CHILD IS THE ABILITY TO SUCCEED AND THRIVE IN LIFE.**

**OUR ONE-TO-ONE MENTORING MODEL IS CLASSIFIED AS A TOP INTERVENTION AND PREVENTION PROGRAM FOR CHILDREN. IT INVOLVES RECRUITING, SCREENING, INTERVIEWING, TRAINING, MATCHING, SUPPORTING AND EVALUATING.**





# **HACK**

**Matching BIGS & LITTLES with a focus on PREDICTING SUCCESS &  
ensuring CHILD SAFETY & INCREASING RETENTION**

# **CODE**

**CREATE AN APP to make the process of SIGNING UP MENTORS EASIER**

**(help with matches + social validation)**

# **FREESTYLE**

**RECOMMEND 'OUT-OF-THE-WORLD' SOLUTIONS TO CURRENT BBBS CHALLENGES**

**(recruiting BIG's, matching them to Littles, reference check BIG's, retention rates etc. - what ideas do you have to recommend)**

**the data**

# Data - overview

1. A data dictionary and other notes are pinned in the data slack channel
2. TREK output is available at <https://hack01.datachambers.com:5601/login>
3. Primary data sources
  - Match details (bsv/parquet)
  - Application interviews (doc/docx)
  - Youth outcome surveys (bsv/parquet)
4. Join keys
  - ChildPartKey – little's numerical id
  - VolPartKey – big's numerical id
  - MatchNum – little-big pair's numerical id
  - InterviewNum - numerical id for the interviews associates with a match
5. Timeline
  - New data: 1/2015 – 2/2017
  - Old data: 1/2009 – 12/2014

# Data – context

## 1. Onboarding Process

- Application to BBBS
- Interview with a MSS
- Match littles to their bigs
- Little takes baseline YOS

## 2. YOS = youth outcome survey

- The first YOS completed sets the baseline scores for that little
- A little will complete their first YOS when they are first matched or at age 9, whichever is later
- A little will complete additional YOS when about once a year after they've been matched and when the match is closed
- Not all matches have a YOS on file

## 3. Successful matches

- A successful community based match is one that lasts at least a year
- A successful site-based match is one that lasts at least a year or the full length of time for the site (full school year, etc.)

## 4. RTBM = ready to be matched

## Data - files

```
/home/shared/bbbs/  
├── general_data_dictionary.bsv  
├── interviews  
├── match_details_data_dictionary.bsv  
├── matches  
│   ├── active  
│   │   ├── match_details_new.bsv  
│   │   └── youth_outcome_reports_new.bsv  
│   ├── all  
│   │   ├── child_volunteer_keys.bsv  
│   │   ├── match_details_new.bsv  
│   │   ├── match_details_old.bsv  
│   │   ├── youth_outcome_reports_new.bsv  
│   │   └── youth_outcome_reports_old.bsv  
│   └── unsuccessful  
│       ├── match_details_new.bsv  
│       ├── match_details_old.bsv  
│       └── youth_outcome_reports_new.bsv  
├── question_ids.bsv  
├── unmatched  
│   └── rtbm_reports_new.bsv  
└── yos_data_dictionary.bsv
```

**tech**



## hardware stack

- 5 node hadoop cluster (hat tip DataChambers)
- Node specs:
  - 32 cores
  - 128GB RAM
  - 6TB storage



## **stack**

- anything that can compile to a .jar (like java, scala, etc.)
- JDBC connection available through Hive
- Python, via pyspark and Anaconda
- Apache Spark

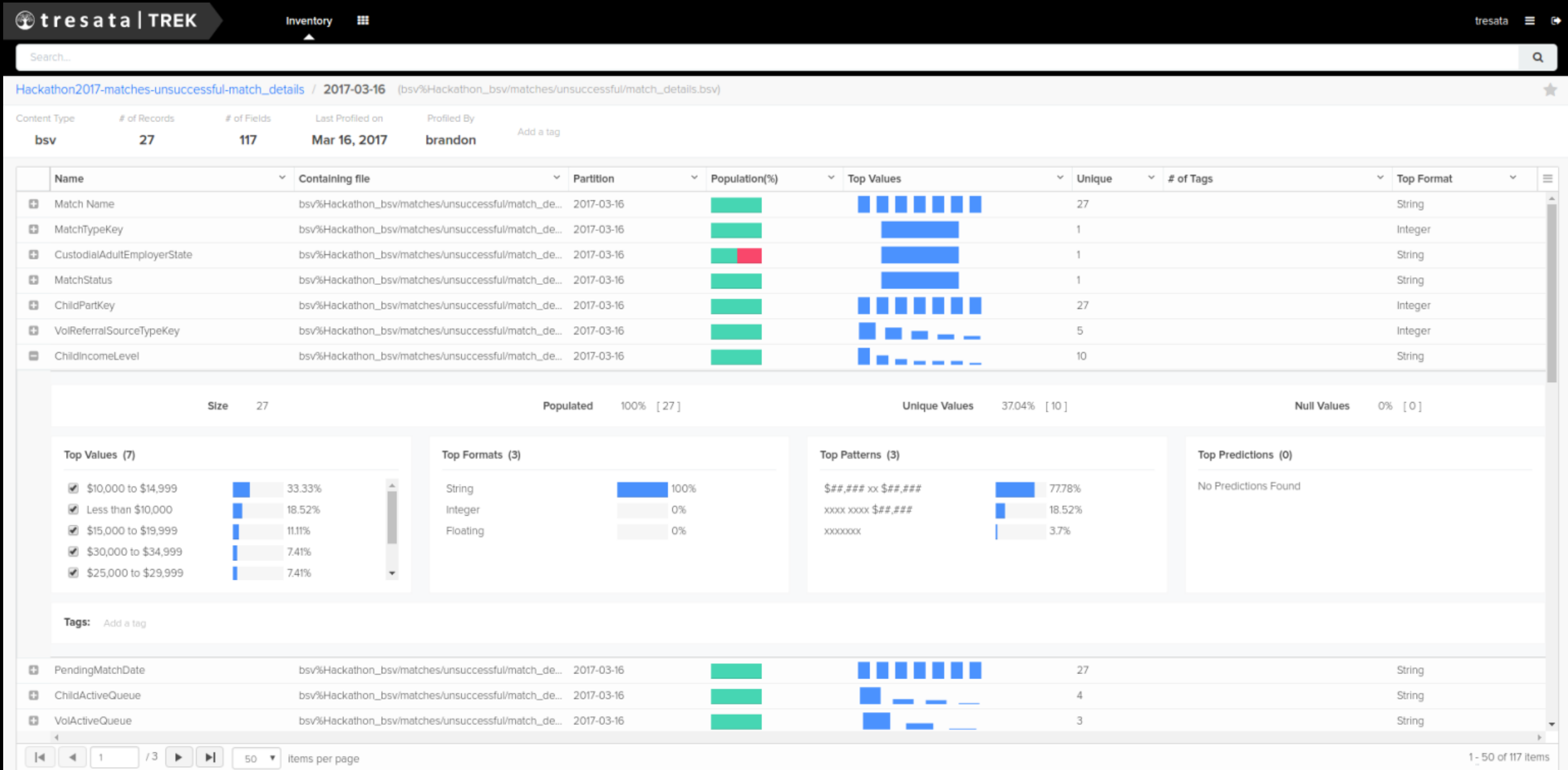
# logistics

- SLACK access
- DO NOT LEAVE WITH ANY DATA
- YES, WE ARE SPOT CHECKING

# Tresata Data GUIDE

- In keeping with our promise, this year we give you access to yet another **awesome kickass** Tresata software
- Tresata's Record Exploration Kit (TREK) has automatically built a **data dictionary**
- This is completely **interactive**
- Allows for **quick, deep data inventory management**
- **Located here** - <https://hack01.datachambers.com:5601/login>
- **HINT:**  
when you are stuck...deep in data muck...  
may this be your periscope...and make your code dope

# Tresata TREK



judging

## **a few hints before you get started**

- understanding of the data (it is complex)
- scale trumps complexity in algo
- Ability to tease out sophisticated insights
- Merge data we haven't given you
- Bottlenecks in the process
- PATTERNS (ANY pattern – checks, herringbone, tie'dye)
- Simple yet powerful algorithms

## how you will share results

- format – your call
- time – 5 minute presentations for shortlisted teams
- Creativity (yours), predictability (machines), scalability (your code) are all important
- judges will be revealed soon... they are tough, but they are awesome



# process

## 2 parts:

1. Shortlisting - starts @ 7am tomorrow

Sign up for appointments via google spreadsheet, link will be on Slack

Panel: Pete Murphy, Kevin Ledford, Mike Keating, Tim Reagan, Dr. Dulin, Koert Kuipers, Abhishek Mehta, Chase Cabanillas, Jayesh Mori, Richard Morris

2. Final Presentations – start @ 12pm

10 Teams – 3 Teams for each category (HACK, CODE, freeSTYLE) + 1 wildcard

# final judges panel



Dr. Tim Chartier,  
Davidson



Donna Dunlap ,  
BBBS



Keva Walton,  
Charlotte Chamber



Abhishek Mehta,  
tresata

**the prizes**

upping the ante for year 5

\$10,000 HACK PRIZE

\$5,000 CODE PRIZE

\$2,000 freeSTYLE PRIZE

**Also new this year...**

**community prize**

ARE YOU READY FOR THIS

## **If you haven't registered yet...**

Make your way to the tech help desk (1<sup>st</sup> floor through the café) to:

1. Get registered
2. Input your team details
3. Get user access to the VPN (one per team)
4. Get your hackathon t-shirt

**questions?**

**[www.hackathonclt.slack.com](http://www.hackathonclt.slack.com)**



# #GOHACKCLT

@charlottehacks | #HACKATHONCLT

<http://www.github.com/tresata/hackathon2017>