

Outcomes Working Group Webinar 14: FINCA - EXPERIENCE

6 October 2016

SPEAKER:

Scott Graham, FINCA

Director of Research and New Business Initiatives



FINCA®



Agenda

- **Introduction**
- **Presentation - FINCA**
- **Discussion with participants**

Our speaker today



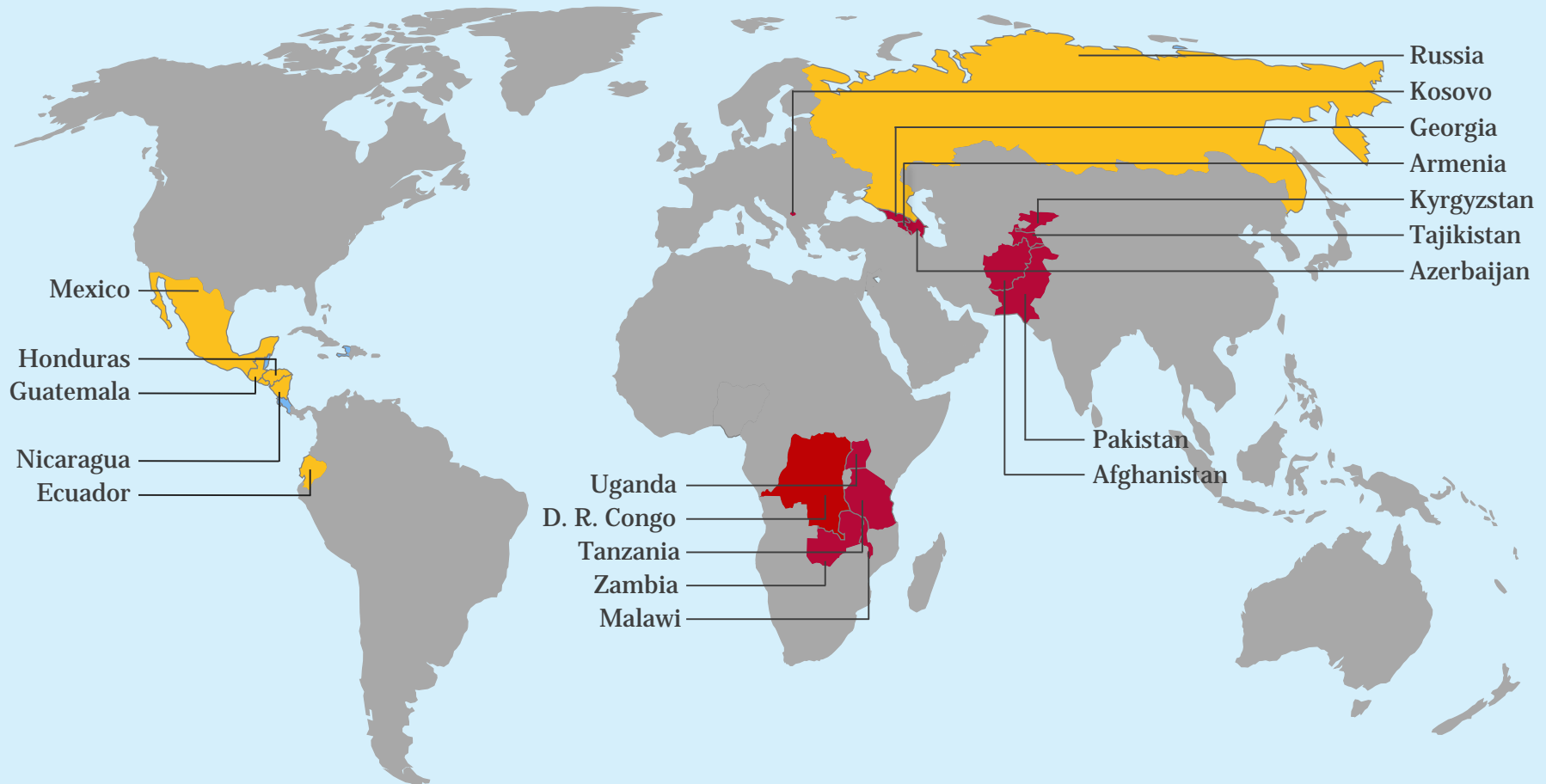
- Scott Graham
- 22 year microfinance career
- Previously Country Director of operations in Malawi and South Africa.
- Led the development of FINCA's Social Performance platform and its customer research practice.

Agenda

- Introduction
- Presentation
- Discussion with participants

FINCA network

5



- HH Composition and Consumption
- Living Standards (Health, Education, Basic Services)
- Business Performance, Seasonality
- Employment and Job Creation
- Women's Empowerment
- Client Aspirations

- Demand and Use of Services
- Client Satisfaction
- Competitive Positioning, Loyalty and Brand Awareness
- Branch-level Performance
- Mapping and Efficiency Analysis

Contents

- **Overall Approach**
- **Measuring Change in Employment**
- **Mission Monitor (Instrument and Output)**
- **Data Quality Pitfalls and How to Avoid Them**

Overall Approach

Household Survey

- **Consumption**
- **Living Standards**
 - Family Composition
 - Employment
 - Education
 - HH Assets

-
- Select **HH variables** to predict **HH income level**

Similar to PPI

every 3-5 years

Enterprise Survey

- **Profitability**
- **Employment**
 - Total Employment
 - % due to FINCA
 - Wage Quality

• **Sector and Other Variables**

-
- Segmentation by **sector of activity** to predict **income and employment generation**

Mission Scorecard

- **1 minute phone survey**
 - Based on the HH and Enterprise variables that are correlated with our key outcomes.
-
- Predictive estimate of **the characteristics of our client population** and our **job creation impact** at a point in time.

quarterly

Measuring Employment Impact (and Attribution)

Control Attributes

- **Sector**
- **Income / Expenses / Profitability**
- **Duration**
- **Growth Expectations**
- **Formality**
- **Current Employment****
 - full-time / part-time
 - paid / unpaid
 - temporary / permanent
 - registered / unregistered
- **Wages Paid****

Changes Observed

- **Change in employment since the most recent FINCA Loan****
- **Attribution to FINCA**

Employment Impact

	EMPLOYMENT PER FIRM		% DUE TO FINCA	
	Baseline		Baseline	
	Avg	Distr.	Avg	Distr.
Trade	5.1	15%	50%	15%
Services	3.2	25%	60%	25%
Manufacturing	8.4	10%	20%	10%
Crops	3.0	35%	55%	35%
Livestock	3.5	15%	40%	15%
Total	3.98	100%	50%	100%

Illustrative data

Average employment per firm is 3.98 people (including owner).

FINCA's loan is responsible for 50% of the employment at this point in time.

Assumption: these ratios will remain the same over a reasonable period of time.

From Snapshot to Moving Picture

	TOTAL EMPLOYMENT PER FIRM					
	Baseline		Q1		Q2	
	Avg	Distr.	Avg	Distr.	Avg	Distr.
Trade	5.1	15%	5.1	30%	5.1	40%
Services	3.2	25%	3.2	15%	3.2	5%
Manufacturing	8.4	10%	8.4	15%	8.4	25%
Crops	3.0	35%	3.0	25%	3.0	20%
Livestock	3.5	15%	3.5	15%	3.5	10%
Total	3.98	100%	4.55	100%	5.25	100%

Total employment per firm is increasing due to growth in outreach to **Trade** and **Manufacturing**

Illustrative data

	% OF EMPLOYMENT DUE TO FINCA					
	Baseline		Q1		Q2	
	Avg	Distr.	Avg	Distr.	Avg	Distr.
Trade	50%	15%	50%	30%	50%	40%
Services	60%	25%	60%	15%	60%	5%
Manufacturing	20%	10%	20%	15%	20%	25%
Crops	55%	35%	55%	25%	55%	20%
Livestock	40%	15%	40%	15%	40%	10%
Total	50%	100%	47%	100%	43%	100%

FINCA's contribution to employment is decreasing because of drop in outreach to **Services** and **Crops**.

Illustrative data

FINCA Mission Monitor

	Base-line	Target	Q2	Q3
Total Result	51%	68%	54%	57%
1 Increased Access to Financial Services by the Poor and Financially Excluded	49.4%	49.5%	45.3%	47.5%
% clients below NPL	7.0%	0.0%	7.4%	9.6%
% clients NPL < 160% NPL	20.7%	21.0%	21.8%	28.4%
% clients NPL < 240% NPL	24.5%	25.0%	25.8%	33.6%
% female clients	56.0%	56.0%	50.4%	52.4%
% rural clients	75.0%	75.0%	31.0%	27.2%
Branches, ATMs and POS per 1000 clients	56.0%	56.0%	109.1%	109.1%
2 Employment, Wages and Income Generation	67.2%	68.0%	68.4%	68.5%
jobs created by FINCA as a share of total jobs	38.9%	39.0%	38.3%	40.0%
wages above minimum wages as a share of total wages	68.7%	70.0%	72.0%	70.9%
enterpr. with income > NPL as a share of total ent.	94.0%	95.0%	94.9%	94.5%
3 Living Standards	50.5%	51.2%	46.3%	47.6%
% HH with no home ownership	36.6%	35.0%	n/a	n/a
% HH without access to Internet	80.0%	80.0%	62.5%	54.8%
% single-parent HH	22.7%	22.7%	26.4%	28.2%
% female headed HH	22.0%	22.0%	35.2%	40.4%
% HH. main earner < secondary education	77.4%	80.0%	60.9%	66.8%

	Base-line	Target	Q2	Q3
4 Empowerment	66.7%	67.0%	60.7%	61.2%
% clients who attain their goals	94.0%	95.0%	86.7%	84.4%
% female-headed enterprises	54.0%	54.0%	48.6%	50.5%
% jobs created by female enterprises	52.0%	52.0%	46.8%	48.7%
5 Responsible Finance	28.6%	92.9%	45.7%	55.7%
A - Sets targets for social performance	0%	100%	50%	75%
B - Monitors social performance regularly	0%	100%	50%	75%
A - SP incorporated into annual operating plans	0%	100%	0%	0%
B - Ensures proper data collection, analysis and reporting	100%	100%	100%	100%
A - Implements reliable customer surveys (Customer Satisfaction, etc.)	100%	100%	100%	100%
B - Makes decisions on product design, services and delivery channels based on client data	0%	100%	20%	40%
1 = self assessment, 2 = in process, 3 = already certified	0%	50%	0%	0%

17 client-level indicators
7 institution-level indicators

Implementation

- HQ designs a standard questionnaire, in consultation with the field.
- HQ digitizes the instrument, installs data quality filters, and trains a subsidiary survey manager.
- HQ and subsidiary staff devise sampling methodology and random selection of respondents.
- Subsidiary staff collect data, following the primary and reserve sampling lists.
- HQ and Survey Manager monitor data quality throughout fieldwork, following-up on questionable answers.
- HQ calculates indicators and prepares analysis with local staff.
- HQ and local staff brief management.

The Sources and Consequences of Bad Data

Enumerator and Respondent Errors

- **Misunderstanding** the question
- **Rushing** through the question or the section
- **Confusing the units of measure** or **time period** in question
- Entering **wrong data**.
- **Falsified** response



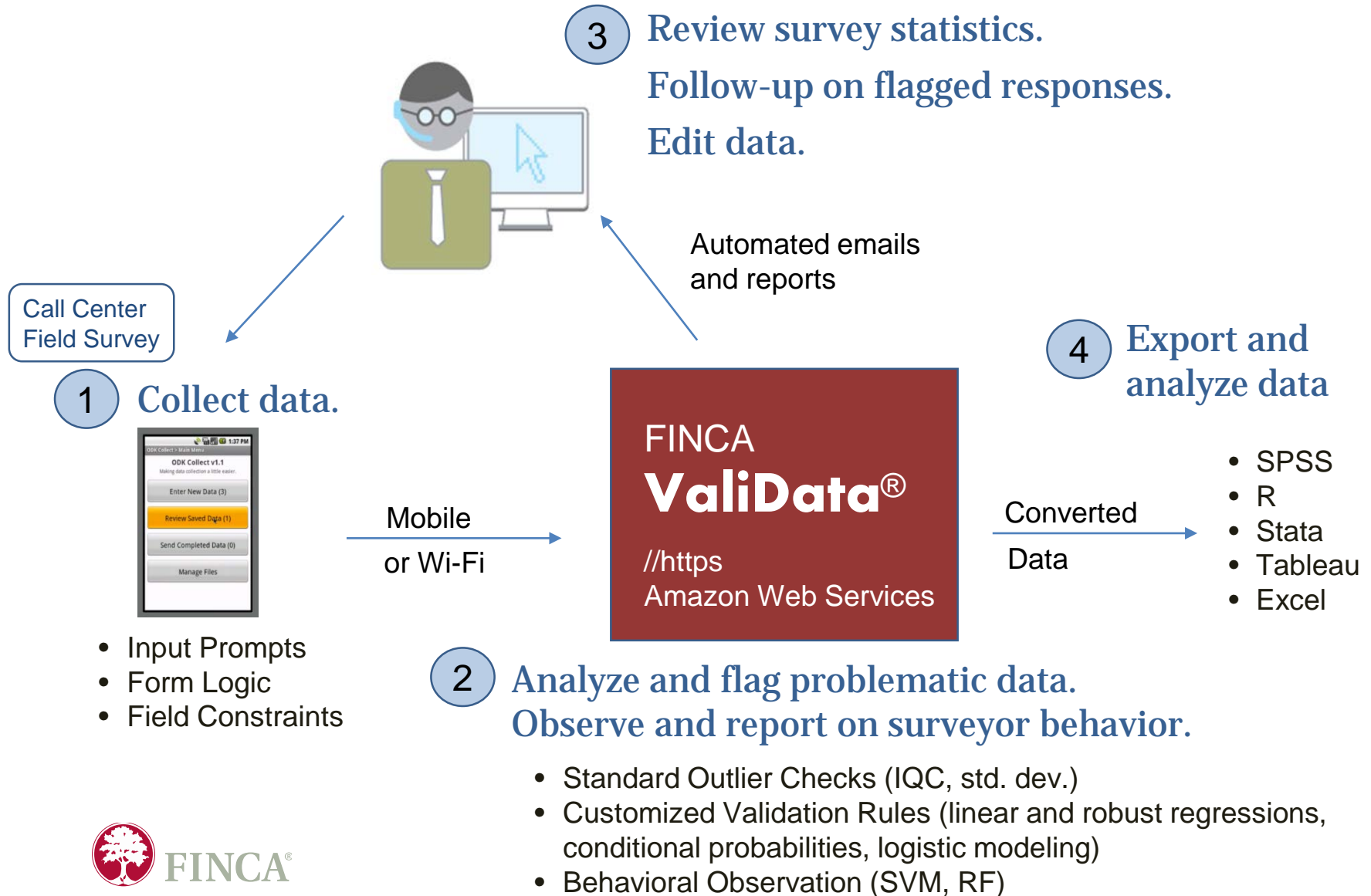
Ex-Post Statistical Remedies

- Data normalization
- Throw out the survey
- Imputations and bootstrapping

Consequences

- Lost **time**
- Wasted **resources**
- Loss of **randomization** and **representativeness**
- Introduction of **bias and noise** in the data
- High **margins of error**, and **confidence intervals**
- **Unreliable** findings, erosion of trust
- **Wrong** predictions

Collecting Clean Data with ValiData



Data Quality Filters

Standard Outlier Checks

- Inter-quartile Checks (IQC)
- Standard deviations

Flag data that fall outside of expected ranges, compared with the responses from other surveys.

Customized Validation Rules

- Linear and Robust Regressions
- Conditional Probabilities
- Logistic Modeling

Ensure that each variable “fits” within the larger data set, considering how it relates to other variables within and between surveys.

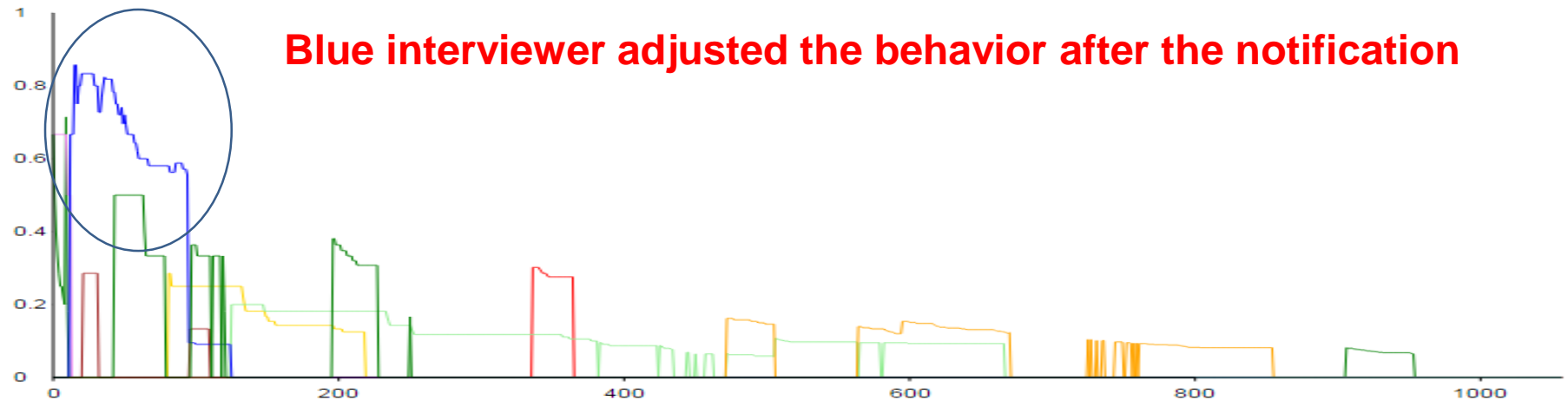
Separation Statistics

- Support Vector Mechanisms (SVM)
- Random Forests (RF)

Identify surveyors whose behavior is anomalous (“separable”), as compared to his past actions and those of his peers.

ValiData An example of separation statistics

Catching **wrong routings** using SVM and Random Forest algorithms, de facto from the first couple of interviews.



- This survey was collected in Pakistan using VALiData.
- From the first couple of interviews ValiData identified highly separable behavior of the “blue” surveyor.
- At first glance on data the “blue” surveyor was interviewing respondents who “didn’t have any other bank accounts except FINCA”
- Survey manager contacted the respondents to check the validity of the surveys
- “Blue” Interviewer was selecting “no account with other financial providers” to skip the whole long section on the competitor’s information. Almost 50% of the respondent he interviewed actual HAD other financial partners.
- The data and the interviewer behavior were corrected immediately.

Concluding observations

- With proper data, a scorecard approach can be applied to other areas of client activity, both in the household and the enterprise, including employment generation.
- With effective controls and cross-validation of data points, it is possible to capture some measures of outcomes at the client level without resorting to a panel study approach (baseline + follow-up)
- The scorecard can be used over time to see how the organization's evolving outreach is driving key mission indicators.
- Technology is an important asset to leverage in order to install good data controls and make the most of limited resources.

Agenda

- Introduction
- Presentation
- Discussion with participants

Thank you

- For follow up, please contact: info@sptf.info, francesinha@edarural.com
- Please note: presentations and recordings from all Outcomes Working Group Meetings are being posted to the SPTF website, working groups page: <http://sptf.info/sp-task-force/working-groups>

APPENDIX

Data Quality Filters

Data Quality Filters

Standard Outlier Checks

- Inter-quartile Checks (IQC's)
- Standard deviations

Flag data that fall outside of expected ranges, compared with the responses from other surveys.

Customized Validation Rules

- Linear and Robust Regressions
- Conditional Probabilities
- Logistic Modeling

Ensure that each variable “fits” within the larger data set, considering how it relates to other variables within and between surveys.

Separation Statistics

- Support Vector Mechanisms (SVM)
- Random Forests (RF)

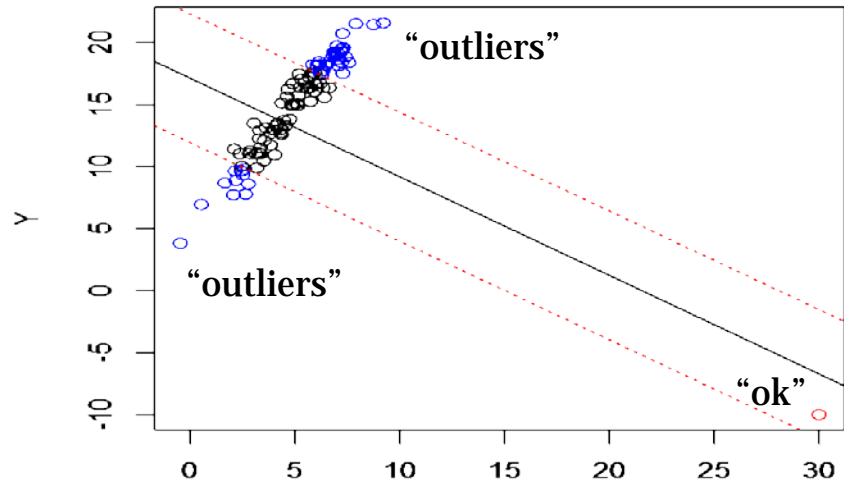
Identify surveyors whose behavior is anomalous (“separable”), as compared to his past actions and those of his peers.

Advanced Outlier Checks

22

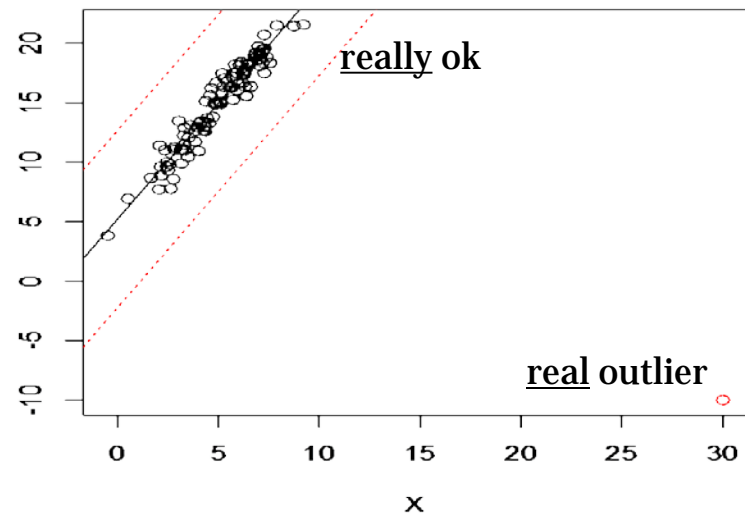
- Simple outlier checks are problematic. With their breakdown point of 0% all it takes is one incorrect observation to turn the entire model in the wrong direction. We end up with type II (red dot) and type I errors (blue dots).

OLS regression skewed by one outlier



- A robust approach uses a combination of inverse weighting on leverage and y-error weighted m-estimation. As can be seen, the approach identified the outlier correctly.

Robust regression using m-estimation



Comparing Outlier Methodologies

	Identified Outliers	Actual Outliers	% correctly identified
IQR	1200	50	4%
St dev from mean/median	1050	50	5%
Linear Regression	100	10	10%
Robust Regression	50	48	96%

FINCA Enterprise Study in Pakistan

n = 900 surveys
50 variables
45,000 data points

per post-
survey
follow-up

Of all the outliers identified by the robust regression approach, 98% are confirmed by the respondents to be actual outliers and the correct values were provided.

ValiData

Catching fake interviews using SVM and Random Forest algorithms, *de facto* from the first couple of interviews.

24

- This survey was collected in El Salvador using ValiData.
- From the first couple of interviews ValiData identified highly separable behavior of the “purple” surveyor.
- Survey manager contacted the respondent to check the validity of the surveys
- Almost 80% of the respondents were not visited.
- Green surveyor was producing highly inaccurate GPS readings due to the lack of knowledge how to capture it. He adjusted the behavior after the notification as can be seen in the picture.

