

# Hit the deck running!

*The Ten Steps to Follow for Your First 100 Days*

**Customer Analytics & Intelligence Conference  
San Francisco  
August 2011**



Jim Porzak – Senior Data Scientist

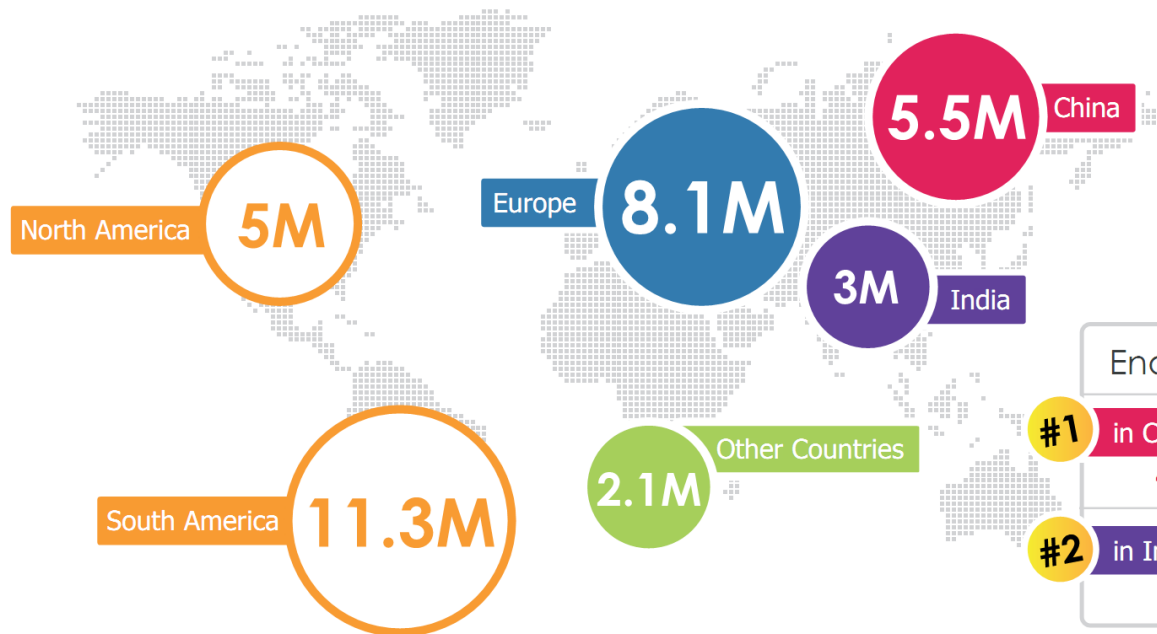
viadeo 

# Who is Viadeo?

## *Professional Networking with a Global Reach*

Global Reach, Local Approach

**35**  
million  
members



2 local brands

Enormous potential in Asia

#1 in China  天際網  
www.tianji.com  
457M Chinese Internet Population

#2 in India  ApnaCircle.com  
52M Indian Internet Population

viadeo 

# The Ten Steps

1. Talk to everyone!
2. Gather data resources.
3. Install *your* tools.
4. Explore the data sets.
5. Define global vocabulary.
6. Cherry pick focus areas.
7. Do selected analyses.
8. Integrate into a story.
9. Draft 6-month road map.
10. Present!

## Underlying Themes for First 100 Days

- KISS
- Crawl, Walk, then Run
- Explore & Discover
  - Not Build!
- These are ongoing efforts.



# 1. Talk to everyone!

## The Execs

- Needs & wants
- Set expectations for first 100 days
- Keep close to your champion

## The Troops

- Where the day-to-day institutional knowledge lives!
- Keep them in loop re goals & progress

## Your CI Customers

- What do they need to succeed?
- Get history, successes, & frustrations
- Current KPI's, if any



## 2. Gather your Data Resources.

### **DBA's are your best friends**

- If not, make it so!
- Show them what you are doing with “their” data
- Food helps!

### **Operational IT**

- Accounting
- Order processing
- Billing
- CRM
- ...

### **Site Developers**

- They know customer behavior
  - Logs & web analytics
  - Internal data base
- Food really helps!

### **Front Line Troops**

- Excel sheets
- Sticky notes



### 3. Install *your* Tools.

*Institutional constraints?*

*(Covered during your hiring interviews!)*



# 3. Install *your* Tools – Jim’s tool kit

## Collaboration

- Wiki – [Google Sites](#)
- Source Code – [SVN](#) or [GitHub](#)

## Database

- [MySQL](#) & derivatives:
  - [InfiniDB](#)
- [Hadoop](#) & friends
  - [Datameer](#)
- [SQLite](#)
- [DbVisualizer](#)
- (ETL ?)

## Analysis & Reporting

- [R](#)
  - [R Studio](#)
  - [RevolutionAnalytics](#)
- [Tableau](#)
- [Jasper Reports](#)

## Utilities

- [TextPad](#)
- [Vedit](#)
- [FireFTP](#)
- [FileZilla](#)
- [7-zip](#)



# 4. Explore Data Sets.

## DBA Theory



## Operational Practice

Are the data elements...

- Reliable?
- Complete?
- Distributed?

Do they have...

- Magic values?
- Context sensitivity?
- Business rule changes over time?

### To Do

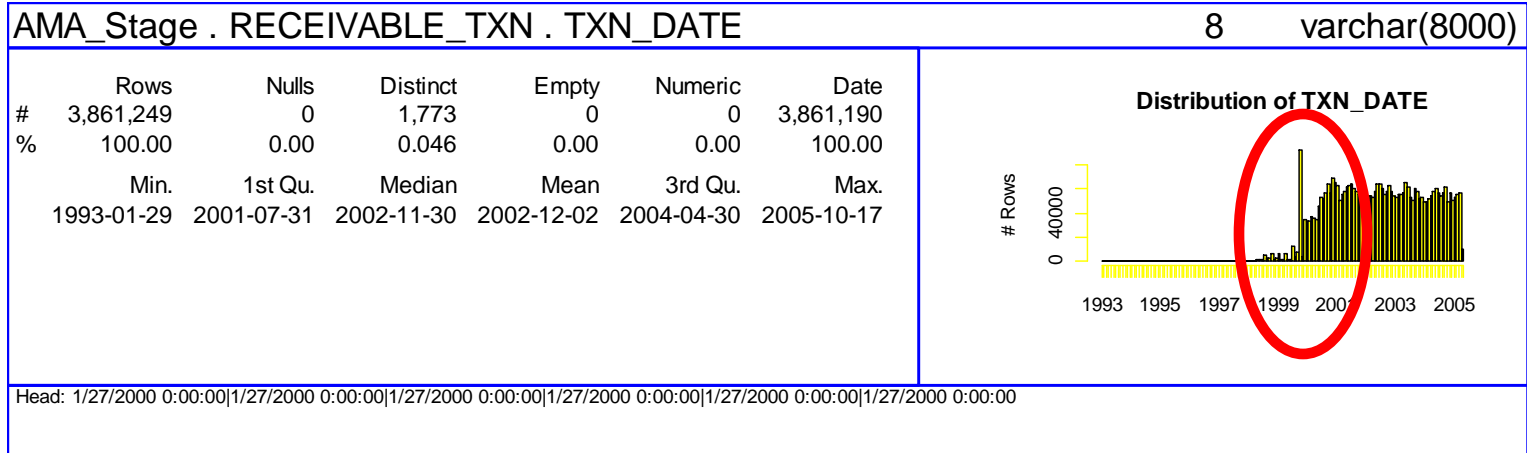
- Profile & QA data sets
- Start Exploratory Data Analysis (EDA)
  - “You can see a lot by just looking.”



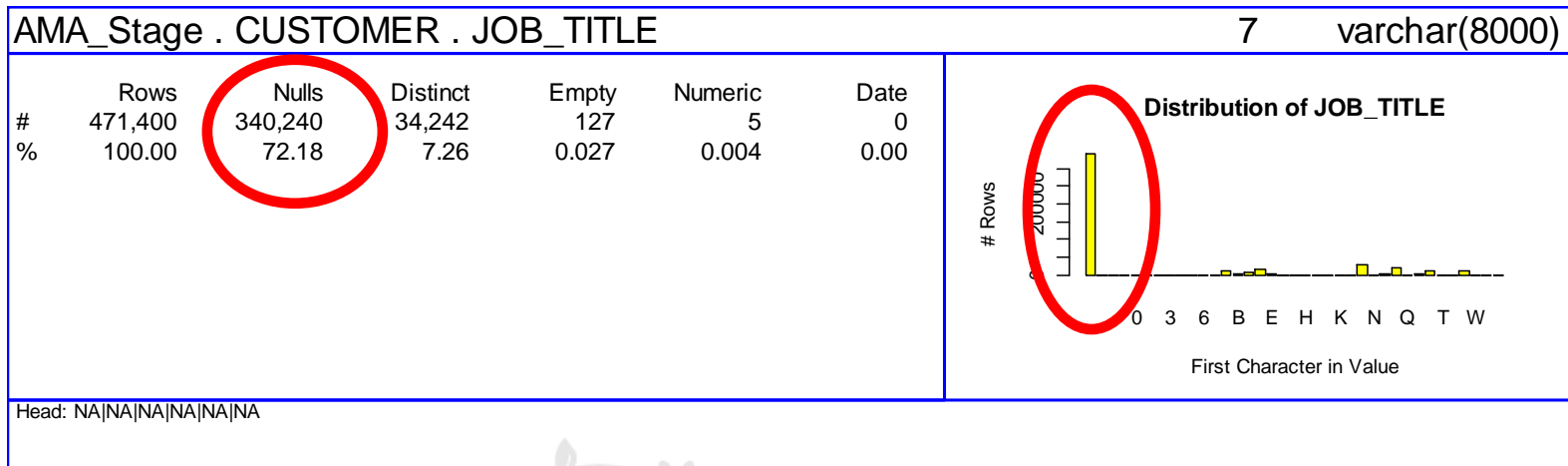


# 4. Explore Data Sets – Data Profiler Examples

## Strange date value



## Sparsely populated value



# 5. Define Global Vocabulary.

Current business terminology is function of each group's focus:

## Marketing

≠ Finance

≠ Product

≠ Operations

≠ Sales

≠ PR

## For Customer Intelligence

- What's a "customer"?
- How many do we have?

## For Subscription Model

- What's a "subscription"?
- How long has a customer subscribed?

Delivering *customer intelligence* requires us to view the world from the customer's perspective.



## 6. Cherry Pick Areas to Focus On.

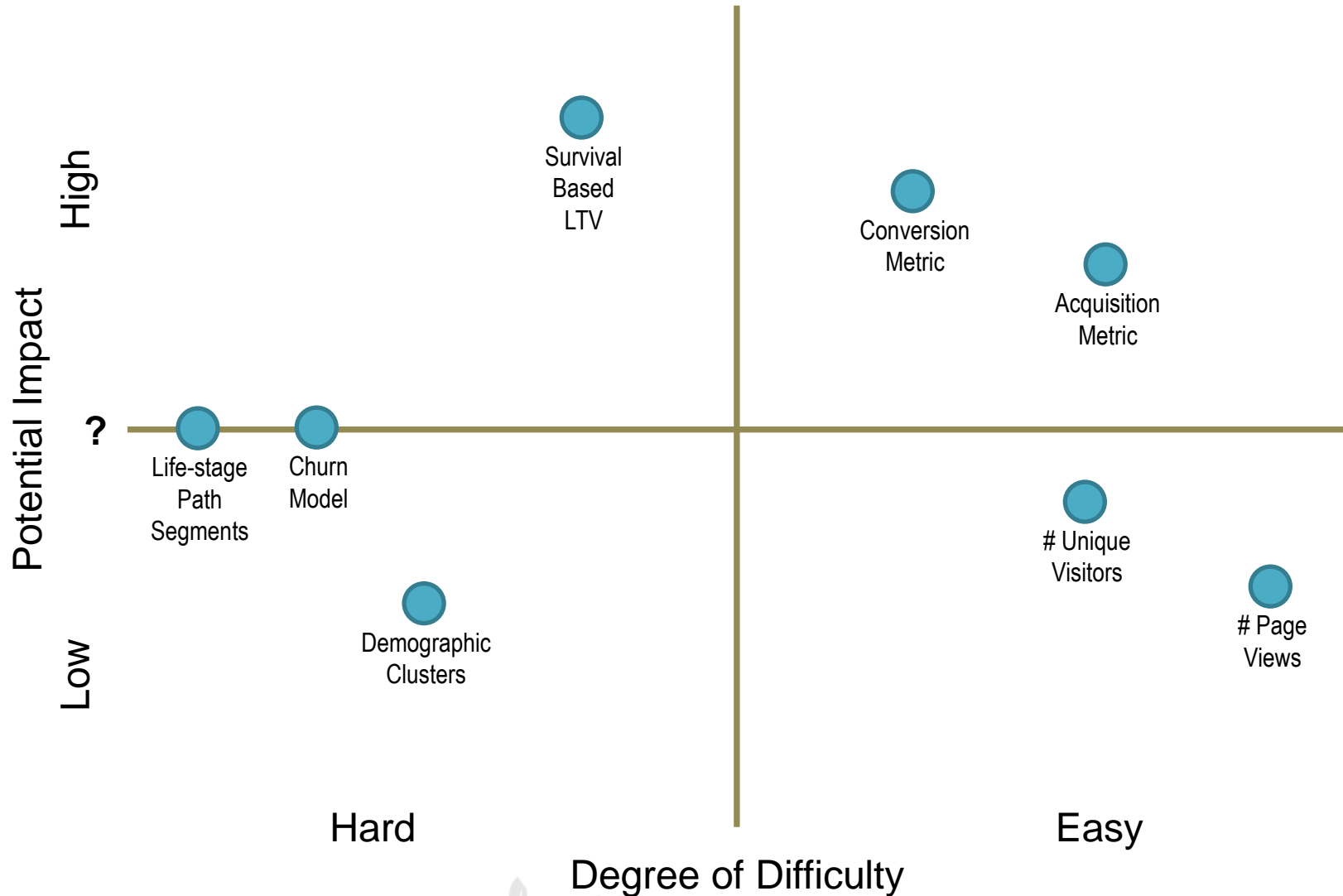
Go for analyses and KPI's with...

- Biggest potential business impact
- Prototype doable in 100 days
- Unique & new concept

How do we sort out priorities?



# 6. Cherry Pick - Using KPI Quadrants



# 7. Do the Selected Analyses.

Remember, we are exploring & discovering, *not* producing.

- Hackers win
- Take a lot of quick looks before deep diving
- Beware of data traps

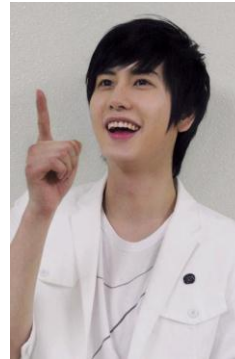
***But ...***

No skimping on:

- Quality visualizations
- QA & cross-checking
- Trending results

**What are the ...**

**Ah Ha's?**



**OMG's?**



viadeo 



## 8. Integrate into a Story.

“If 80% of what you present they already know, there’s a good chance they will believe the 20% that’s new.”

- Know your goal
- Position as partial & preliminary insights
- Pick the best KPI(s) and show
  - Strategic Value
  - Tactical relevance
  - What if?
    - Some % change = some \$ increase



# 9. Draft your 6-month Road Map.

The first 100 days is all about exploration, discovery, & testing concepts.

The next 6 months will be about delivering an initial platform for a selected few KPI's and enabling end user exploration and analysis in a few selected areas.

## Benefits to Organization

- Who will get what?
- Potential business impacts:
  - Strategic
  - Tactical
  - Savings
  - Gains

## Budget you Need

- \$'s
- Resources from other teams
  - This will be harder to get!



# 10. Present



"Before you go any further, let me reiterate that I, for one, see nothing wrong with killing the messenger."



# 10. Present

A short, focused deck

- Selling *one* take-away
- But, with extensive appendix to back up results & explain any new concepts in detail.
- Hints
  - Shop your ideas before the BIG presentation.
  - Evangelize new concepts as you go along.
  - Be sure the techies are on board with new methods



# Review & Discussion

1. Talk to everyone!
2. Gather data resources.
3. Install *your* tools.
4. Explore the data sets.
5. Define global vocabulary.
6. Cherry pick focus areas.
7. Do selected analyses.
8. Integrate into a story.
9. Draft 6-month road map.
10. Present!



And, ping me with  
questions, comments:  
[JPorzak@ViadeoTeam.com](mailto:JPorzak@ViadeoTeam.com)