细分客户Customer Segment:

1.大众，对各种数据感兴趣的人

2.投资者，根据市场数据决定投资策略的人

3.企业决策者，根据各种数据制定企业策略的人

4.数据拥有者，需要有一个平台来可视化自己的数据的人或分享给别人

5.通过对数据的分析来牟利的专业人士

6.更多...

价值主张Vale Proposition:

1.全面，准确，及时的数据推送和搜索服务（数据仓库）

2.美观，多样的数据呈现方式（显示工具）

3.科学，创新的数据分析方法（分析工具）

4.未来也可以考虑提供数据采集服务(大数据中心)

渠道沟通Channel:

1.Mobile App

2.Web App和网站

3.线下工具

核心资源Key Resources

1.数据平台

2.可视化工具

客户关系：Customer Relationships

1.兴趣小组，提供用户之间交流的平台

2.专题讲座培训，如何使用数据

3.积分机制，成就机制

收入来源：revenue stream

1.广告2.VIP功能 3.特殊数据收集

核心业务Key Activities

1.数据获取模块开发

2.终端模块开发

重要合作key partnerships

数据提供商

成本结构Cost structures

人力成本

设备成本

宽带成本

收入：

广告收入

VIP包月包年收入

咨询收入

seease.com

viszee.com

jusuu.com

数据抓取：

已处理：

世界银行  
[http://data.worldbank.org.cn/](http://data.worldbank.org.cn/" \t "_blank)

未处理：

[http://finance.sina.com.cn/mac/#](http://finance.sina.com.cn/mac/)

<http://calendar.hexun.com/global433_1.shtml>

<http://finance.stockstar.com/finance/macrodata/jijinkaihu.htm>

<http://www.cnnic.net.cn/>

<http://visualizing.org/data/browse>

http://news.qq.com/bigdata/

http://www.qianzhan.com/qzdata/

http://www.stats.gov.cn/tjsj/

http://data.baidu.com/index.html

http://datacenter.cngrain.com/  
  
http://www.askci.com/data/  
各国数据  
http://www.gapminder.org/data/  
搜房数据  
http://fdc.soufun.com/data/

<http://www.ebrun.com/data/>

元数据：最基本的数据单位(根据分类来创建不同的collection)

-数据结构[  id（自动生成），数据名称DataKeywords ]

-数据值[  区域，具体周期，主体1(Target1)，主体2(Target2)，数据时间，来源，数值s，更新时间，备注s，是否为预测值 ]

组合数据：数据选择器，用来选择系列数据

-组合条件s：区域条件，周期条件，主体1条件，主体2条件，数值条件，外部数据

-组合数据s:

-组合信息：id，名称，star，follow，comment，点击率，所属目录s，文字说明

-可选显示：图表s，地图，说明

数据类别：用来对数据进行分类

类别标识

类别名称

上级类别s

区域：用来描述数据所属地理区域范围

区域ID

区域识别号

区域名称

上级区域

主体：数据的描述对象

-名称（）

-类型（行业，机构，用户，公司，国家，城市，私人等）

-类型值

-标识ID

-备注

数据主体：（显示用）  
数据唯一识别号(GUID)

数据名称

数据一级目录（可多选）

数据二级目录（可多选）  
数据地理范围（全球，国家，城市，区域，公司，私人）

数据产业范围（全社会，第一产业，第二产业，第三产业，轻工业，重工业，居民生活）  
数据周期（年，季度，月，日，不定期，定期）  
数据来源（组织，机构，网站，私人）  
数据单元  
数据保存（内部，外部）

数据管理员

数据内容（时间，数值，备注）

{ID:"",Name:"",Range:"", Period:"",Src:"",Unit:"", Store:"",ExtLink:"",Data：["1956-1-1":12,"1957-1-5":15]}

**前端技术**

跨平台，包括web, android, ios, phoneGap做成hybrid app

**框架**

UI框架：JQuery Mobile

**图表显示**

**d3.js**

支持大多数浏览器，不支持的如IE6可通过 Node.js后端生成SVG文件，使用svgweb显示

**后端架构**

数据抓取Scrapy(Python)+MongoDB+raids

**需求驱动开发**

**第一阶段，数据显示，基本的SN**

V1.0.0：数据显示功能

  分类显示数据

  数据使用多种图表

  人工采集数据

V1.0.1: SN功能

  基本的用户注册，评论

  Star/Follow

  更多方式的排序，包括按时间，按人气

v1.02:数据比较功能

  多种数据比较

**第二阶段：数据的采集，显示，管理，完善SN**

v1.1.0

  互联网数据检索，人工干预

  用户自定义数据发布

V1.1.1

  数据分享到XXX

  用户纠错

v1.1.2

  大数据融合

**第三阶段，框架化，外部资源整合，SN细化**

  提供API给其他网站

AK(Abstract Keyword)，抽象关键字，主要指数据信息中会出现的高频词，如同比

CK(Concrete Keyword)，具象关键字，主要指数据名称，如高考人数

数据采集

Data crawler设计

1. 分两类，数据集中式网站和数据分散式网站

Tasks:

-多语言表

Objectid,name, Chinese:AAA,English:BBB

- AK management : add/update/remove

抽象关键字管理页面：

Objectid,loc\_name,

- CK management: add/update/remove

具象关键字管理页面：

Objectid,loc\_name

- Catalog management: add/update/remove, level 1, level 2

- Country management : add/update/remove