

## 数读互联网消费时代的大学生金钱观

### 摘 要

随着互联网和移动互联设备的发展，数据的来源已经被极大的拓宽，但是如何从海量的数据中获取价值，挖掘信息，则是一个挑战与机遇。设计作为一门融合了多种思维方式和创新手法的学科，可以给这个问题提供一个解决思路，那就是数据可视化。

本次毕业设计的主题和数据来源是互联网消费行为，因为互联网消费已经深深的影响了我们的生活方式，一方面给我们带来了便利和实惠，大大的减少了交易的门槛和成本，但是另一方面，网上购物也产生了很多问题。

我们意图通过网页的和数据可视化的方式，从多个角度用收集来的数据呈现网络消费的利弊；也在不同年龄，性别的行为可视化比较之中对现阶段大学生对金钱观有一个初步的认识。在可视化的过程中，我们希望能保持一个开放的态度，不同的人可以得到不同的结论和启示，这也正是数据可视化的魅力。

**关键词：**数据，可视化，网络消费，观念，网页

## Visualizing Consumption on the Internet and Attitudes towards Money of the Students

### ABSTRACT

With the development of Internet and mobile devices, data have been increasingly accessible in the last decades. But how to get insights and values from the huge amount of data is still a big challenge and opportunity for us. Design, as a subject combined multiple thinking and innovating methods, could give us a new perspective to deal with the challenge, which is data visualization.

The data and theme of this capstone project both come from the consumption behaviour on the Internet, because Internet consumption have affect us in many different ways. On the one hand, it make trades more free and convenient; on the other hand, it brings us many troubles as well.

We intent to build a website to visualise this phenomena from different perspectives and show a initial picture of students' attitudes towards money. We hope everyone keep open-minded when they browse this visualisation and get insights and conclusions of his/her own.

**Key words:** data, visualization, consumption, Internet

## 目 录

1 引 言.....	4
1.1 大数据时代下的机遇.....	4
1.2 数据可视化的意义.....	4
1.3 主题选择.....	4
2 可视化理论部分.....	5
2.2 信息可视化设计框架.....	5
3 方案设计.....	6
3.1 双十一数据可视化.....	6
3.2 消费群体数据可视化.....	6
3.3 金融数据可视.....	6
3.4 借贷数据可视化.....	6
4 技术实现.....	7
4.1 数据分析技术.....	7
4.2 网页搭建技术.....	7
4.3 数据可视化技术.....	7
5 结论和展望.....	9
5.1 结论.....	9
5.2 展望.....	9
参考文献.....	10
谢 辞.....	11

## 1 引言

### 1.1 大数据时代下的机遇

大数据是未来社会发展的一个重要趋势。随着社会的发展进步，越来越多的电子设备，智能设备进入我们的生活，随之而来的则是数据在数量与质量上的急剧飞跃，而且可以预见随着新技术的产生与应用，数据将会变得越来越普遍与可信。数据代表着权力，掌握数据资源的组织或者公司会在竞争中占有优势。

然而原始数据并不会产生任何价值，只有对数据进行处理才能从数据的矿山中发掘出价值。对数据进行处理发掘价值的方法有很多，但是我们作为设计系的学生，在这里提供一个更加感性的方案，就是用数据可视化的方式来展示数据的魅力与价值。

### 1.2 数据可视化的意义

数据可视化相对于其他技术类型的数据分析方法，有如下好处：直观。根据生理学研究显示，视觉是人类信息输入最重要的一部分，视觉的信息输入带宽是第二位触觉的十倍，是听觉和嗅觉的一百多倍。所以，用可视化的形式来展现数据，可以最大程度上被人脑理解和接受。可拓展性好。数据可以从不同角度，不同方面进行分析，而数据可视化提供了最大的可能性。不同的数据可视化逻辑和可视化方法可以展现出同一个数据的不同方面和不同内容，可以展示出数据更深层次的内容。

### 1.3 主题选择

在调研分析的基础上，我选择了互联网消费作为可视化分析的主题和对象，原因如下：数据可得性好。由于互联网消费的蓬勃发展，互联网电商已经积累了大量的数据资源，一些电商愿意将这些数据资源进行有限度的公开，这就为数据可视化分析提供了可能性。

贴近生活。互联网消费并不是离我们很遥远的事情，而是实实在在影响我们生活的事情。几乎我们每天都要从互联网上购买商品，进行消费。只有这样的话题才更加贴近我们的生活，和用户产生一种共鸣。

争议性。互联网消费在带给我们巨大便利的同时，还有很多问题乃至危机。例如互联网消费催生了消费贷款，助长了过度消费的风气，很多人由于网上消费入不敷出，严重影响了社会的秩序，造成了很多负面影响，也引起社会对很大争论。而数据可视化可以对主题又一个深层次分析，解构其中对因果。

## 2 可视化理论部分

### 2.1 信息可视化的宏观过程

根据乔治 A 米勒（George A. Miller）的研究，人类短期记忆一般一次只能记住 5-9 个事物。这一事实经常被用来作为限制导航菜单选项到 7 个的论据；然而关于神奇的“7,加 2 或者减 2”还是引起了激烈的讨论。所以,在信息爆炸时代,怎么在同类的海量信息中,让受众,更加容易认同自个儿的,就得尽量不违反“7±2 原则”。

7+2 原则，是认知科学中最基础的定理之一。制作一个可视化图表，至少包含以下环节：收集—整理—设计—输出。然而数据的收集和整理则是重中之重。每次可视化都是为了解决特定问题的,所以,面对海量以标准形式收集的数据,要事先思考怎么针对领域问题合理抽取对应的数据。罗列出要解决的问题，什么数据对于可视化的需求有用？这需要数据分析者有一个强有力的准确方向。重要的是,所有问题,都要归结为一个单一形式来表现。

### 2.2 信息可视化设计框架

信息图形化设计，在明确想要展现的内容后，脑子里面有个基本的框架和第一时间想到的展示形式，但这样就足够了么？当然还要运用视觉元素的造型，色彩的选取，动态等赋予图表更好的视觉体验。如果用代码形式在网页上面做出相应的布局，信息图形就不仅仅是一个简单的静态页面了，而是一个可以活动的大型数据流的输出展示。例如 GE 做的“美国历史生育数据做的可视化”。可以看到美国历史上各个婴儿出生的波峰波谷与当年的一些重大事件、发明和政策间的关系。通过鼠标在时间轴上的位移，数据的视觉效果随之变化，动态的看出美国在不同时期的剩余数据的变化。

在确定表现形式，非人工实现设计，而是要通过数字工具统计分析后输出图形时。就需要设计者设计合理的输出合成的形式。例如下图

面对日益丰富的可视化数据分析工具。一些相关工作人员则设计出不同的展现形式，来帮助用户更好的进行视觉风格的选择。例如 datavlab 公司开发的插件可以根据你提供的数据帮你选择适合的展现形式，例如：treemap，tree，stream，scatterplotMatrix，force，matrix，bubble，chord。设计师在选择表现形式的同时一定要考虑到受众是否能够很好的去接受和吸收信息，设计师需要了解视觉心理对用户浏览时的影响。伴随着互联网的发展，对于优良设计的要求也越来越高了。

信息（数据）可视化是一个较为复杂的领域，绝不是美化图表那么简单，最终目的不是可视，而是从可视化中进行数据挖掘，信息传播，帮助分析问题。

## 3 方案设计

### 3.1 双十一数据可视化

双十一的数据的特点是数据量巨大，双十一这一天的交易数据远远超过其他天的数据；所以在数据分析的基础上，对原始数据进行了类比分析。我们用 2000 个方块的类比来编码从 5.11 到 11.11 这大半年的所有店铺每天的销售数据；相应的，我们用颜色来编码时间序列，用动态效果来表达各个数据之间的逻辑关系。而且动效让不同数据之间产生了一个交互对比的过程，更加直观的可以看到双十一的销售量的数据特点。

网页的流水瀑布式布局让数据可视化的动效有了充分的空间，也很自然的成为叙事逻辑的一部分，在用户不断向下拉伸网页的这个过程，也是一个顺序浏览可视化效果的过程。

### 3.2 消费群体数据可视化

消费群体的数据特点是数据的层次很多，维度不同，不同维度的数据叠加可以得到数据从不同角度分析的效果。而我们采用的是一种层层叠加的数据分析方法，由简单到复杂的来叙述消费群体的不同特质。

分析的角度有：性别维度，年龄维度，收藏维度，购买维度和时间维度，每个维度都有对应的数据可视化逻辑进行分析；在最后有一个 Detail on Demand 的数据集，用户可以根据自己的状态和兴趣来选择不同的数据角度。

### 3.3 金融数据可视化

金融数据的主题是“二八定律”。金融数据在人数和钱数上的不匹配是这部分可视化逻辑的核心。大多数的底层人民占据的财富却极少，而人数上最少的富人则占有了社会的绝大多数财富。

在这个基础之上又增加了地理纬度的信息，在金额和地理上都存在着非常明显的马太效应或者二八定律，这让我们联想到科幻小说北京折叠，在这种可视化的结论之下就更加能够引起我们的反思。

### 3.4 借贷数据可视化

网络借贷是互联网消费的一个负面影响，在全民网购的大背景之下，也有很多人以为网购背上债务而因此踏上不归路。数据的维度多样，数据的类型多样，包括时间数据，类型数据，空间数据等等。而数据可视化的策略就是用不同的图形编码方式来解决不同类型数据的不兼容问题来达到可视化上的协调性与统一性。

这是作品最后希望能够引起反思的部分；数据来源源于网络新闻的互联网公开数据，在数据的基础之上对不同数据类型和数据维度进行综合编码可视化，得到一个数年间恶性网贷事件的全景图。

## 4 技术实现

### 4.1 数据分析技术

中国的大数据应用处在起步阶段，但随着电信和银行领域开始对大数据技术和服务产生浓厚的兴趣，未来三年我国大数据市场将突破百亿元。2012 年市场规模达到 4.7 亿元，2013 年增至 11.2 亿元，增长率高达 138%，2014 年，保持了与 2013 年基本持平的增速，增长率为 114.38%，市场规模达到 24.1 亿元，未来三年内有望突破 150 亿元，2016 年有望达到 180 亿规模，其中增长率最高的是存储市场，将达到 60.8%，服务器市场的增长率则是 38.3%，远远高于其他非大数据产品相关的市场，由此来看整个行业发展空间巨大。

进入 2014 年以来，大数据受到各界广泛关注，已渗透到金融、医疗、消费、电力、制造以及几乎各个行业，大数据的新产品、新技术、新服务也正在不断地涌现。我国大数据产业集聚发展效应开始显现，出现京津冀区域、长三角地区、珠三角地区和中西部 4 个集聚发展区，各具发展特色。北京依托中关村在信息产业的领先优势，快速集聚和培养了一批大数据企业，继而迅速将集聚势能扩散到津冀地区，形成京津冀大数据走廊格局。长三角地区城市将大数据与当地智慧城市、云计算发展紧密结合，使大数据既有支撑又有的放矢，吸引了大批大数据企业。珠三角地区在产业管理和应用发展等方面率先垂范，对企业扶持力度大，集聚效应明显。大数据产业链上下游企业合作意愿强烈，各集聚区间的合作步伐加快，产学研协同创新发展初见成效。数据可视中国的大数据应用处在起步阶段，但随着电信和银行领域开始对大数据技术和服务产生浓厚的兴趣，未来三年我国大数据市场将突破百亿元。2012 年市场规模达到 4.7 亿元，2013 年增至 11.2 亿元，增长率高达 138%，2014 年，保持了与 2013 年基本持平的增速，增长率为 114.38%，市场规模达到 24.1 亿元，未来三年内有望突破 150 亿元，2016 年有望达到 180 亿规模，其中增长率最高的是存储市场，将达到 60.8%，服务器市场的增长率则是 38.3%，远远高于其他非大数据产品相关的市场，由此来看整个行业发展空间巨大。进入 2014 年以来，大数据受到各界广泛关注，已渗透到金融、医疗、消费、电力、制造以及几乎各个行业，大数据的新产品、新技术、新服务也正在不断地涌现。

我国大数据产业集聚发展效应开始显现，出现京津冀区域、长三角地区、珠三角地区和中西部个集聚发展区，各具发展特色。北京依托中关村在信息产业的领先优势，快速集聚和培养了一批大数据企业，继而迅速将集聚势能扩散到津冀地区，形成京津冀大数据走廊格局。长三角地区城市将大数据与当地智慧城市、云计算发展紧密结合，使大数据既有支撑又有的放矢，吸引了大批大数据企业。珠三角地区在产业管理和应用发展等方面率先垂范，对企业扶持力度大，集聚效应明显。大数据产业链上下游企业合作意愿强烈，各集聚区间的合作步伐加快，产学研协同创新发展初见成效。化中，前期的数据挖掘和数据分析虽然没有在最后的结果中直接呈现，却是整个设计环节中最重要最基础的部分。

具体到技术方面，此次数据的爬取分析的程序是 Python 并且利用了相应的第三方库进行辅助；在数据爬取方面使用的是 BeautifulSoup 的第三方库；而数据的分析储存则是基于 Pandas 库来进行的。

## 4.2 网页搭建技术

网页搭建会有很多技术方式，这一次我使用的是 Bootstrap 的网页框架。Bootstrap 是目前主流的 CSS 框架，先前也有提到，在起初我是很不喜欢 Bootstrap 的框架，认为他抢走我大部分的工作且缺乏特色，但在研究后会发现他并非我想的那么缺乏特色，完全看开发者的使用及熟练度

我们也可以看到网路上有相当多的 HTML Template 都做的非常美，但其实底层都是 Bootstrap，有些已经调整得相当不像是原有的 Bootstrap 样式。所以在熟练的情况下，Bootstrap 也是能做出超乎水平的网页版型。

对于熟悉 CSS 的网页设计师来说，到会觉得自己就能解决这样的状况，所以 Bootstrap 在起初就不认为那么的必要，但对于大型网站、系统来说，自己写 CSS 也会产生几个问题：

持续性的维护：架构是否良好将影响 CSS 是否好维护套件的使用：再引用其他第三方套件时，是否能够依据需求调整 CSS 符合风格。

## 4.3 数据可视化技术

这次我选用的数据可视化技术是基于 JavaScript 的 P5.js 编程库。跟严肃的 C/C++，学究的 MatLab/R 或者神经兮兮的 Python 相比，Processing 是一种很文艺的语言，比如会有常量叫 HALF\_PI 和 TWO\_PI，你也不用纠结太多细节，比如传值还是传址，有很多让喜欢用 Processing 写代码，类似画个速写，Processing 的程序名默认也是 sketch\_xxx。Processing 是基于 Java 的，P5.js 是 Processing 的一个衍生版本，完全用 Javascript 实现。

由于是基于 javascript 实现，所以 P5.js 所写的就是可以在浏览器里直接运行的。如果你用的是“现代”的浏览器，比如 Chrome、firefox 甚至 Safari，哪怕是移动版的，都是可以运行的。你做出来的东西可以方便的展现和传播，从传播角度来说比 matlab、python 什么的方便多了。

对于数据可视化来说，P5.js 也提供了很多友好的函数和数据接口，可以方便而快速的进行数据可视化的实验性探索和数据可视化的网页搭建。



## 5 结论和展望

### 5.1 结论

作品通过对于互联网消费数据行为的分析和可视化呈现，向观众展示了一个数据视角下的互联网消费时代的各个层面，在展览过程中有很多观众产生了共鸣。展示的过程中有很多观众表达了对于数据可视化与网页的兴趣。

### 5.2 展望

1. 对于数据的进一步获取，展示更多的数据细节和数据维度。由于时间紧张很多数据维度挖掘的并不够深，在数据表达的基础上，开放性的结论还不足以更加深入细致的表达主题。所以在未来课题还将继续，从不同的方面挖掘数据，表达数据，则会是一个阶段的重点。

2. 更加丰富的交互方式，充分展现“Big Picture - Zoom in - Detail”的逻辑。

## 参考文献

- [1] 唐. 诺曼 设计心理学 2: 如何管理复杂[M] 中信出版社, 2011
- [2] 唐. 诺曼 情感化设计[M] 电子工业出版社, 2005
- [3] Nikki Knox Jun How to use persona Empathy mapping[M] , 27 2014
- [4] David McCandless Collins Information is Beautiful: The Information Atlas[M] (2010-02)
- [5] Edward R. Tufte, Envisioning Information[M], Graphics Pr, ( 1990-12)
- [6] Donald A. Norman, Tamara Dunaeff, Things That Make Us Smart[M], HarperCollins UK, (2013-10)
- [7] Manuel Lima, 视觉繁美 - 信息可视化方法与案例解析 [M] 机械工业出版社 2013-8-1
- [8] Tamara Munzner Visualization Analysis & Design [M] A K Peters/CRC Press, 2014-12-3
- [9] Julie Steele / Noah Iliinsky 数据可视化之美 [M] 机械工业出版社 2011-6
- [10] 肖勇 名师设计实验课: 信息设计 湖北美术出版社, 2010
- [11] Manuel Lima, The Book of Trees. Visualizing Branches of Knowledge
- [12] Manuel Lima, Visual Complexity: Mapping Patterns of Information
- [13] Robert Bringhurst, The Elements of Typographic Style
- [14] Tufte, The Visual Display of Quantitative Information

## 谢 辞

首先我要感谢我的导师曹静老师，感谢您对我整个毕业设计的耐心指导与包容。您认真负责的态度带给了我很大的帮助和鼓励，助我最终顺利完成毕设。感谢曾给予我教诲和帮助的老师：吴洁、忻颖、孙效华、杜钦、许扬波、曹静等等。感谢你们一直以来对我的帮助和教导！同时我还要在此感谢我的朋友和同学们，感谢你们一直以来的帮助。这次毕业设计的过程中也受到了许多同学的帮助和支持，对我的方案有很大启发。

我还要感谢一直以来默默支持我的家人们，感谢父母为我完成学业而辛劳地付出，愿你们平安健康。

装

订

线