# Title this and that

Process Book

Alex Mocanu, Manana Lortkipanidze

Data Visualization, EPFL 2018

Introduction

This process book serves as a description of our project for data visualization class at EPFL, Fall 2018. Starting from initial abstract drafts, sketches and ideas to concrete final product. Explaining on the way reasonings behind every analysis or design step. This process book guides its readers through our thought process from the day first, providing along details regarding functionality of visualization, design choices, its implementation and usage.

At the end of this process book we discuss possible usage scenarios, impacts and future improvements of our visualizations. Furthermore, we provide technical details for people interested in developing our work.

Overview

Wikipedia is not only the most widely used online encyclopedia, it is also considered as one of the most visited/popular websites in the world. Us and our generation probably remember times, when one had to scroll through hundreds of pages of heavy books, just to find definition/information regarding one of the aspects of interest. Wikipedia makes everyone’s life easier while being completely free. In our opinion this deserves appreciation.

Motivation

Our motivation was to study human behaviour based on Wikipedia data set. As already mentioned above, Wikipedia is a free and openly editable Encyclopaedia. It is created by volunteers and is one of the most popular websites. Therefore, political news, climatic changes, new trends, technologies, movies and basically all-important life events impact its content, edit requests, visitor counts and frequencies. We think it would be interesting to have an insight into human mindset and behaviour patterns trough Wikipedia data.

Visualization tries to answer question: can we identify major events based just on sudden peaks in number of visitors for a specific page? Which pages get affected for specific category of event? How does increase in visitors on pages affect visitor count of its hyperlinked pages? More specifically, how far does the interest spread in terms of related pages? How long does the interest last depending on type of events?

Furthermore, we could separately study categories and their ranking. Specifically, we are mainly interested in studying structure of pages and identifying the most popular ones. We will pose the following questions: Which pages are the most controversial?

Target Audience

Our target audience is anyone interested in human mindset and reflection of real-life events and trends on human behavior regarding Wikipedia.

Inspiration

Concept

Dataset

Designs

Initial Idea

Implementation Model

Technical details

Implementation

Technical details