

# CityBike Visualisation

Dalma Bencsik\*  
1103075

Dagcan Mermi†  
1126575

Thomas Nußbaumer‡  
0806012

Iris Reihs§  
0401380

## ABSTRACT

Unser Projekt soll als ein Tool für die Mitarbeiter der Firma CityBike dienen. Die von uns entwickelte Visualisierung ermöglicht verschiedene Grafiken darzustellen, die grundlegende Informationen für Analyse der Stationen in unterschiedlichen Aspekten anbieten. Neben einer allgemeinen Darstellung hat man mithilfe unseres Tools die Möglichkeit, die Entlehn- und Rückgabedaten von den Stationen zu analysieren, oder sogar zwei beliebige Stationen miteinander zu vergleichen, wo der Benutzer interaktiv auswählen kann, über welchen Aspekt der Daten er mehr Informationen erfahren möchte.

**Keywords:** CityBike, Radfahren, Visualisierung, Verkehrsfluss.

## 1 MOTIVATION

Als Projektthema im Rahmen der Lehrveranstaltung Visualisierung haben wir uns die Visualisierung für die Firma CityBike ausgesucht. Da die Firma selber schon mehrere Möglichkeiten für die Analyse der Daten anbietet bzw. auch in der Lehrveranstaltung haben sich mehrere Teams mit dem selben Thema beschäftigt, wollten wir das Thema unter einem anderen Aspekt annähern.

### 1.1 Benutzer

Ein wichtiger Punkt, worin sich unsere Visualisierung von anderen Datenanalysen unterscheidet, sind die Benutzer. Unser Projekt soll den Mitarbeitern der Firma bei der Analyse der Stations- bzw. Verkehrsdaten helfen. Mit diesem Tool wird es für sie ermöglicht, früher vorhersagen zu können, zu welchen Zeitpunkten die Boxen höchstwahrscheinlich leer sind und zu welchen Stationen sie noch zusätzliche Fahrräder transportieren sollen.

### 1.2 Tasks

In der globalen Ansicht können die Mitarbeiter allgemeine Informationen (wie die Nationalität der Kunden oder die Schadenshäufigkeit der Bauteile) gleich nach dem Öffnen des Tools erhalten. Diese Informationen können der Firma bei der Kostenrechnung helfen, bzw. grundlegende Daten für die Optimierung der Reparaturkosten liefern. Außerdem haben sie mithilfe der Nationalitätsdaten die Möglichkeit, die Spracheingabe bei den Terminals an die Bedürfnisse der Kunden anzupassen. Die zwei weiteren Ansichten ermöglichen eine ausführliche Analyse der einzelnen Stationen. Bei der Ansicht Verkehrsfluss können sie die Entlehn- und Rückgabedaten interaktiv analysieren, während Detailansicht noch zusätzliche Möglichkeiten für die Analyse und den Vergleich von Stationen in unterschiedlichen Aspekten anbietet, natürlich nach den Wünschen des Benutzers.

\*e-mail: a1103075@unet.univie.ac.at

†e-mail: a1126575@unet.univie.ac.at

‡e-mail: a0806012@unet.univie.ac.at

§e-mail: a0401380@unet.univie.ac.at

## 1.3 Daten

Die Daten für die Visualisierung haben wir von der Firma CityBike [3] bekommen. Jedoch hatten wir zuerst Schwierigkeiten mit der Zusammenführung von zwei verschiedenen Datensätzen, da diese unterschiedliche Schlüsselfelder für die Stationsbezeichnung hatten. Deswegen mussten wir mit der Firma CityBike in Kontakt treten, und gleich danach haben wir einen zusätzlichen Datensatz bekommen, welcher mit den Fahrtdaten verknüpft werden konnte und die weitere Arbeit ermöglicht hat.

## 2 VERWANDTE ARBEITEN

Im Rahmen eines anderen Projektes wird schon eine Möglichkeit gezeigt, wie man die Verkehrsdaten von Radfahrer sinnvoll visualisieren kann [1]. In diesem Video werden Fahrten in New York, zwischen dem Central Park und der Brooklyn Bridge gezeigt. Dabei wurden nicht nur die Routen eingezeichnet, man kann auch sehen zu welcher Uhrzeit eine Fahrt stattfand.

Außerdem bietet die Firma CityBike Wien auch selbst schon auf ihrer Webseite eine Analyse für die Fahrten an [2]. Hier werden in einem Barchart Fahrten, angemeldete User bzw. mithilfe eines Liniendiagramms gefahrene Kilometer von 2012 und 2013 verglichen. Unser Projekt versucht aber mehr Interaktivität in die Visualisierungen einzubauen, und sich bei den Analysen nicht nur auf die Fahrtdaten zu beschränken, sondern auch mehrere Aspekte für die Visualisierung anzubieten.

## 3 ANSATZ

Unsere ersten Gedanken für eine sinnvolle Visualisierung waren, dass wir den Verkehrsfluss zwischen den Stationen und den Gebrauch der einzelnen Stationen darstellen, indem wir die Wienkarte in den Vordergrund stellen und die Informationen auf dieser Karte anzeigen.

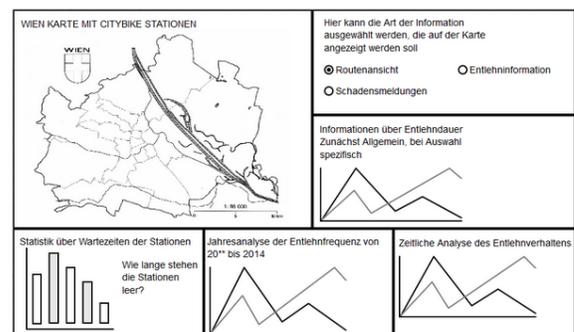


Figure 1: Unsere ersten Ideen.

Für die Analyse der Routen wollten wir den Verkehrsfluss mithilfe von Pfeilen darstellen, wo die Dicke der Pfeile auch Informationen darüber bieten hätte sollen, wie häufig diese Routen benutzt werden. Hier hatten wir noch vor, auch andere Visualisierungen, wie zum Beispiel die Wartezeiten der einzelnen Stationen mittels Balkendiagramm zu implementieren.

Da wir nach der Präsentation dieses ersten Lo-Fi-Prototypen das Feedback bekamen, dass so eine Darstellung auf der Karte für einen

Mitarbeiter eher unrelevant ist, weil die Leute, die bei dieser Firma arbeiten, kennen sich meistens mit der Lage der Stationen schon aus, entschieden wir uns, die ganze Darstellung nochmal durchzudenken und die Visualisierungen anders zu gestalten.

### 3.1 Hi-Fi-Prototyp

Bei der Implementierung unseres Hi-Fi-Prototypen veränderten wir unser Grundkonzept für die Visualisierung der Stations- und Verkehrsdaten, und statt die Wienkarte in den Vordergrund zu stellen, dachten wir uns eine andere Art aus, wie wir die Daten am sinnvollsten visualisieren könnten. Jedoch warfen wir die Idee mit der Kartendarstellung nicht komplett weg, aber nach dem neuen Konzept sollte sie eher nur zusätzliche Information anbieten, als ein Hauptelement für die Visualisierung dienen. Die Routendarstellung war aber weiterhin ein Problem für uns, weil auch wenn wir diese Implementierung schaffen konnten, war die Darstellung ziemlich unübersichtlich und es waren keine Linien mehr erkennbar.

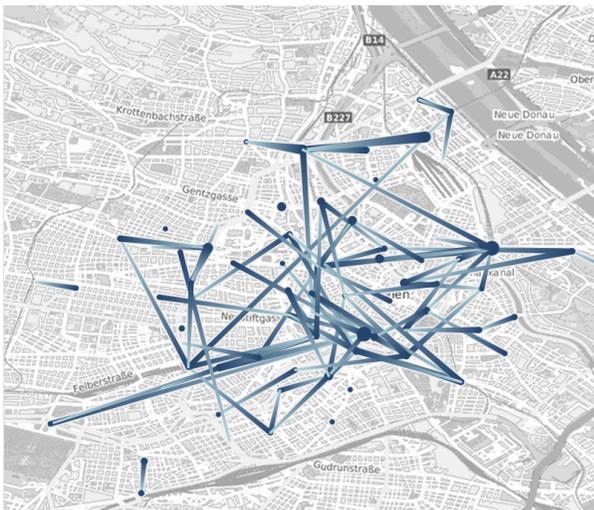


Figure 2: Probleme bei der Routendarstellung.

Deswegen entschieden wir uns so, diese Art der Darstellung einfach wegzulassen, und uns eher an die Entlehn- und Rückgabedaten, Auslastungen und andere stationsspezifische Daten zu konzentrieren. In unserem neuen Prototypen erstellten wir zwei Bubble Charts für die Visualisierung der Entlehnzentren bzw. der Rückgabezentren, welche Diagramme miteinander auch verknüpft wurden, das heißt, zum Beispiel bei der Auswahl einer Entlehnstation wurden im anderen Bubble Chart die Rückgabestationen dementsprechend angezeigt. Außerdem wurden die Entlehn- und Rückgabefrequenzen für die jeweilige ausgewählte Station mithilfe vom Liniendiagramm dargestellt, und zusätzlich auch die Auslastung dieser Station. Bei der Auslastung hatte man sogar die Möglichkeit, nach verschiedenen Wochentagen zu filtern, jedoch funktionierte diese Funktion bei den anderen Diagrammen noch nicht einwandfrei. In zwei Balkendiagrammen visualisierten wir noch die Benutzungszwecke der jeweiligen Station bzw. ob die meisten Kunden, die bei dieser Station ein Fahrrad entlehnt haben, eine Touristenkarte besitzen. Als zusätzliche Information boten wir auch eine Darstellung der durchschnittlichen Fahrtdauer für die Fahrten an, welche Fahrten von dieser Station aus gestartet wurden.

Da wir für dieses Konzept positive Feedbacks erhielten, wollten wir die Hauptelemente für die endgültige Version auf jeden Fall behalten. Die Darstellung mit Bubble Charts war aber für die meisten Leute schon ziemlich verwirrend, deshalb suchten wir eine andere

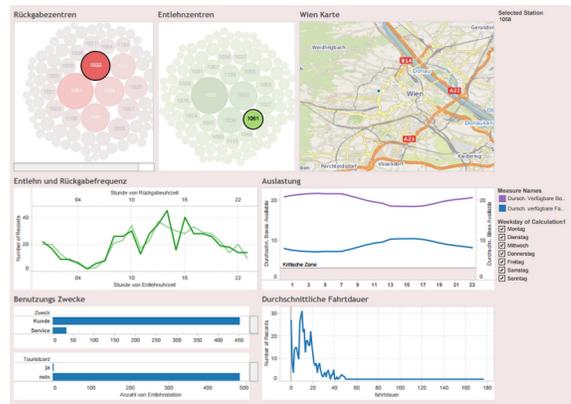


Figure 3: Unser Hi-Fi-Prototyp.

Möglichkeit aus, wie wir die Entlehn- und Rückgabedaten visualisieren, und eventuell auch neue Funktionen einbauen könnten. Um das realisieren zu können, entschieden wir uns für die Darstellung von verschiedenen Ansichten, wobei auch die früher noch nicht funktionierenden Filterungen korrigiert wurden und auch neue Aspekte implementiert wurden. Diese endgültige Version wird im Kapitel Ergebnisse noch ausführlicher diskutiert.

## 4 IMPLEMENTIERUNG

Unser Projekt wurde in Tableau 8.3 implementiert [4]. In den Anfangsphasen haben wir uns größtenteils auf die Datensätze konzentriert da diese zuerst noch bereinigt werden mussten. Im Anschluss darauf hat das Team angefangen die in den Mockups vorkommenden Grafiken in Tableau zu implementieren. Einige der Visualisierungen, wie zum Beispiel die Anzeige der durchschnittlichen Auslastung einer Station, waren vergleichsweise simple zu implementieren. Für andere Visualisierungen waren aus den vorhandenen Daten abgeleitete Kennzahlen notwendig. Diese Funktionalität ist in Tableau mit wenig Aufwand erreichbar.

### 4.1 Kartenansicht

Zunächst war es unser Plan eine Wienkarte in die Visualisierung einzubauen, welches die gefahrenen Routen anzeigen sollte. Dafür gibt es eine eingebaute Funktion welches sich Paths nennt. Um diese Funktion nutzen zu können müssen die Daten jedoch in einer vorgeschriebenen Form vorliegen. Somit wandten wir uns der Aufgabe automatisch die vorhandenen Daten in die gewünschte Form zu bringen. Die erste Schwierigkeit war hier die Zusammenführung zweier Datensätze. Zum einen ging es um die Fahrtdaten und zum anderen um die Stationsdaten. Da die uns zur Verfügung gestellten Stationsdaten ein anderes Schlüsselfeld für die Stationsbezeichnung hatten als die Fahrtdaten, mussten wir mit City-Bike Vienna kontakt aufnehmen. Sie konnten uns dann einen Datensatz zukommen lassen, welche mit den Fahrtdaten Joinbar waren. Mit diesem Datensatz konnten wir eine Visualisierung auf der Wienkarte bewerkstelligen, jedoch waren die so erstellten Pfade zu zahlreich um einen Nutzen daraus ziehen zu können. Um die Kartenansicht zu Bereinigen mussten wir die dargestellten Daten stark reduzieren. Dadurch ging aber die Interaktivität mit der Karte verloren, da wir genauere Informationen über einzelne Fahrten wegschmeißen mussten. Dadurch ging aber der Sinn unserer Visualisierung verloren, da nun nur noch sehr minimalistisch damit Interagiert werden konnte. Um diese Funktionalität einbauen zu können, wäre es notwendig gewesen ein kleines Programm zwischen dem Datensatz und Tableau einzubauen, damit je nach Wunsch des Benutzers die für die Visualisierung notwendigen Daten

automatisch generiert werden können. Das war jedoch mit Tableau nicht möglich wodurch wir im Endeffekt die Kartenansicht nur zur Lagedarstellung der Stationen verwendet haben. Die Darstellung findet mit Hilfe der Maps Funktion von Tableau statt. Durch einen WMS(Web Map Service) werden anhand der angegebenen Geo-Daten die entsprechenden Karten angezeigt. Wir haben dafür den Server von OpenStreetMap verwendet [5].

### 4.2 Interaktionen

Die Tatsache, dass die von uns verwendeten Daten auf drei verschiedene Datensätze verteilt sind, welche sich nicht sinnvoll Joinen lassen, stellte sich als ein Hindernis bei der Implementierung der Interaktivität heraus. Es gibt keine von Tableau aus eingebauten Funktionen zur datensatzübergreifenden Filterung oder Suche. Da jede Grafik allein von einem Datensatz seine Daten beziehen kann, wird es schwierig bei Dashboards mit Grafiken mit verschiedenen Datensätzen eine einheitliche Interaktion zu ermöglichen. Vieles der Probleme konnten wir umgehen, indem wir genügend globale Parameter definierten, anhand dessen die Auswahl des Benutzers erfasst und auf verschiedene Grafiken übertragen werden konnte. Ein Problem damit war die schwierige Wartbarkeit dieser Lösung. Änderten wir die Grafiken und bauten neue Grafiken ein, mussten auch die globalen Parameter angepasst werden. Hier war die Tableau Nutzergemeinschaft sehr hilfreich. Da Tableau ein weit verbreitetes Tool ist, waren wir nicht die ersten die auf diese Probleme stießen. Wir konnten durch frühere Diskussionen in den Nutzerforen viele nützliche Tipps und eher komplizierte Vorgänge in Erfahrung bringen. So konnten wir auch die Visualisierungsauswahl implementieren welches durch ein Drop-Down Feld dem Nutzer es erlaubt, die gewünschte Visualisierung anzuzeigen.

Bei der Implementierung der Entlehn- und Rückgabe Grafik, griffen wir zudem auf Dashboard-Actions zurück. Dadurch wird es ermöglicht, bei Auswahl des Nutzers bestimmte Filter auf die Datensätze anzuwenden.

### 4.3 Story Ansicht

Eines der größten Kritikpunkte aus der zweiten Präsentation war die mangelnde Usability. Um dieses Problem anzugehen verwarfen wir unsere bisherige Gesamtansicht und überlegten uns ein neue Art der Navigation im Tool. Wir entschieden uns für die Story Ansicht in Tableau. Diese Ansicht ermöglicht uns die Navigation zwischen verschiedenen Dashboards, wodurch wir für jede Funktionalität ein eigenständiges Dashboard einbauen können.

Citybike Data



Figure 4: Tabs der Story-Ansicht.

Im ersten Feld "Allgemeine Ansicht" werden dem Benutzer verschiedene stationsunabhängige Daten präsentiert. Diese sind nicht unbedingt hilfreich für unsere Usecases, bieten aber einen guten Einstieg in den Datensatz und präsentieren Daten, die in dieser Form nicht weiter vorkommen.

Im zweiten Feld "Fahrverhalten" befindet sich die Übersicht über die Entlehn- und Rückgabeverhalten der einzelnen Stationen. Der Nutzer hat hier die Möglichkeit nach Top X Einträgen zu sortieren. Zudem kann der Datensatz nach Wochentagen gefiltert werden. Zudem bietet diese Ansicht eine Interaktionsfunktion. Wird auf eine der beiden Diagramme eine Station ausgewählt, so passt sich das andere Diagramm an. Genauer beschrieben werden nach einer Auswahl NUR noch die Daten angezeigt, die auf

die ausgewählte Station zurückgeführt werden. Wird zum Beispiel die Station 1022 im Entlehndiagramm ausgewählt, so werden im Rückgabediagramm nur noch die Daten jener Fahrten angezeigt, die in der Station 1022 begonnen haben. So hat der Benutzer die Möglichkeit stationsspezifische Daten angezeigt zu bekommen. Ein Nachteil hier ist die Auswahl. Um den interaktiven Vorgang sinnvoll weiterstufen zu können, muss der Benutzer die Auswahl zuerst aufheben. Sonst werden Filter verkettet und die angezeigten Daten sind nicht mehr aussagekräftig.

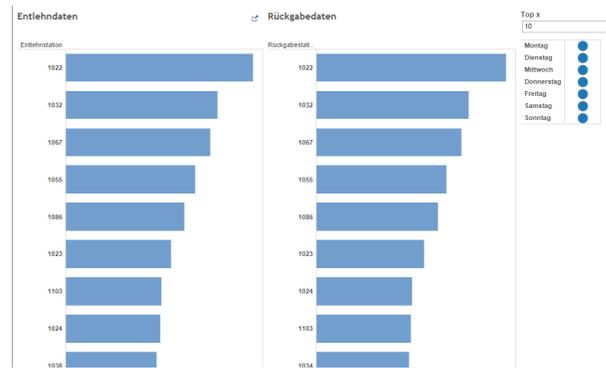


Figure 5: Balkendiagramme der Fahrtenansicht.

Im dritten Feld "Detailansicht" hat der Benutzer die Möglichkeit Daten zu vergleichen. Das macht er indem er im ersten Feld eine Station auswählt und danach die gewünschte Visualisierung wählt. Die entsprechende Vis wird nun angezeigt. Damit können Stationsdaten verglichen oder zwei Datensätze von einer Station auf Korrelationen überprüft werden. Hier hat man ebenfalls die Möglichkeit auf die Filterung nach Wochentagen. Zudem können mehrere Wochentage gleichzeitig ausgewählt werden indem die STRG Taste bei der Auswahl gedrückt wird. Ganz unten sieht man zudem die Position der ausgewählten Stationen auf der Wienkarte.

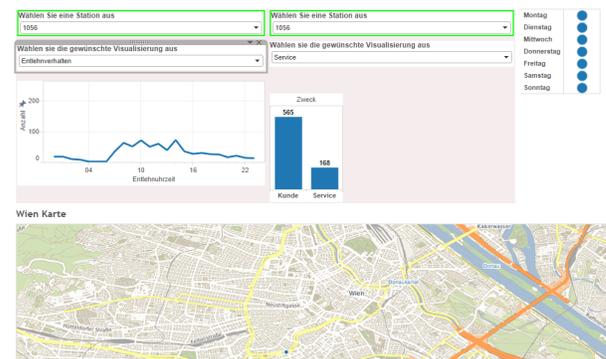


Figure 6: Detailansicht mit Wienkarte.

### 4.4 Bugs

In der Ansicht "Verkehrsfluss" gibt es keine Möglichkeit eine bestimmte Station schnell zu finden falls diese nicht unter den Top Stationen ist. Um diesen Fehler zu umgehen haben wir die Ansichten Verkehrsfluss und Detailansicht mit einem globalen Parameter verbunden. Wird in Detailsicht eine Station ausgewählt, taucht diese Station auf jeden Fall in den Balkendiagrammen der Verkehrsansicht auf und wird zudem farblich hervorgehoben. Diese Lösung fanden wir als optimal für die Problemstellung. Jedoch stellte sich heraus, dass durch eine Implementierungsentscheidung

des Tableau-Teams alle globalen Parameter in der Story Ansicht instanziiert werden. Das heißt für uns, dass unser Lösungsweg zwar in den Dashboards funktioniert, aber in der Storyansicht nicht. Es gibt leider keinen Workaround für dieses Programmverhalten da das Tableau-Team es als "Working as Intended" sieht...

## 5 ERGEBNISSE

Bei unserer Endpräsentation sind wir durch Feedback zu dem Ergebnis gekommen, dass die Variante der Story Ansicht, genauso wenig übersichtlich ist, wie die zuvor verwendeten Bubble Charts. Denn die Story Ansicht ist zwar klarer gestaltet als die Bubble Charts, aber durch das hin und her klicken, wurde auch mit dieser Variation die Usability nicht gesteigert. Wir hätten uns auf ein Fenster konzentrieren sollen, damit der User alles auf einem Blick hat, da sich die Aufteilung in ausschließlich einem Fenster als schwierig erwies, wurde schlussendlich auf die Story Ansicht ausgewichen.

## 6 DISKUSSION

Ein großer Vorteil von unserer Visualisierung ist, dass man eine große Anzahl an verschiedenen Visualisierungsmöglichkeiten hat, was wir durch die Auswahl von unterschiedlichen Ansichten realisiert haben. Mit dieser Einordnung der verschiedenen Aspekte konnten wir unsere Darstellungen relativ übersichtlich halten, und mithilfe des globalen Parameters haben wir es auch geschafft, Filterungsmöglichkeiten für den Benutzer zu bieten, ohne dass diese ausgewählten Filterungen in den anderen Ansichten verloren gehen. Jedoch haben diese Ansichten die Darstellungen voneinander ziemlich getrennt. Außerdem kann es vielleicht auf den ersten Blick ein bisschen verwirrend sein, dass die Diagramme bei der Ansicht Verkehrsfluss die Achsen und die Größen der Balken nicht enthalten. Des Weiteren ist die Möglichkeit für Interaktion bei der globalen Ansicht reduziert.

Durch dieses Projekt haben wir selbst erfahren, welche Schwierigkeiten bei der Erstellung einer Visualisierung auftreten können. Gleich bei der Aufbereitung der Datensätze können schon Probleme auftauchen, und auch wenn man diese gelöst hat, muss man bei jeder Phase neue Herausforderungen bewältigen. Diese Phasen brauchen viel Zeit und Arbeit, bis man eine fertige Visualisierung aufweisen kann.

## REFERENCES

- [1] J. Ferzoco. Citi Bike Rides: September 17th & 18th, 2013. <http://vimeo.com/89305412>, 2013. [Abgerufen: 2015-06-29].
- [2] G. W. mbH. Citybike Wien - Nutzungsdaten. <http://www.gewista.at/DE/Presse/Pressemeldungen/2014/140131/Bildergalerie.aspx>, 2013. [Abgerufen: 2015-06-29].
- [3] G. W. mbH. CityBike Wien. <http://www.citybikewien.at/>, 2015. [Abgerufen: 2015-06-29].
- [4] T. Software. Tableau Desktop. <http://www.tableau.com/de-de/products/desktop>, 2015. [Abgerufen: 2015-06-28].
- [5] terrestris GmbH. Terrestris WMS. <http://ows.terrestris.de/dienste.html#wms>, 2015. [Abgerufen: 2015-06-28].