

Bachelorarbeit

Analyse der Jahresleihdaten des Bike-Sharing-Systems BicikeLJ (Ljubljana, Slowenien)

Clemens Frey

Datum: 26.03.2018

Kurzfassung

Das Bike-Sharing-System BicikeLJ in der slowenischen Hauptstadt Ljubljana wurde in dieser Bachelorarbeit auf das Verhalten der Nutzer analysiert. Die Basis hierfür bildeten die vom Betreiber zur Verfügung gestellten Ausleih- und Rückgabedaten im Zeitraum von August 2014 bis August 2015. Die gelieferten Informationen beinhalteten die Entnahmestation und die Rückgabestelle jeweils mit Datum und Uhrzeit. Ebenfalls vorhanden waren die Nummern der entlehnten Fahrräder, welche allerdings nicht weiterverwendet wurden. Die monatlich getrennten Datensätze wurden aggregiert, in eine einheitliche Form gebracht und anschließend analysiert. Aus den Stationsinformationen wurden die Luftdistanzen zwischen den Verleihstellen errechnet. Aus den angegebenen Uhrzeiten konnten Ausleihdauern und weiters anhand der Distanzen die Punkt-zu-Punkt-Reisegeschwindigkeiten berechnet werden. Durch die Angaben der Entnahme- und Rückgabestellen bei jeder Fahrt konnte eruiert werden, wie stark die einzelnen Terminals im Beobachtungszeitraum frequentiert waren. Die verschiedenen Ergebnisse wurden mittels Diagrammen und Tabellen in Microsoft Excel und durch das Programm QGIS mittels Karten auf Basis der OpenStreetMap dargestellt.

1 Einleitung

Aufgrund der Klimaveränderung und des steigenden Umweltbewusstseins gewinnen klimaneutrale Fortbewegungsmittel für den Individualverkehr immer mehr an Popularität. Die Anzahl der Städte, in denen Bike-Sharing-Systeme angeboten werden, wächst kontinuierlich. In manchen Ballungsräumen gibt es bereits mehrere Anbieter, da die Zielgruppe aktiver und umweltbewusster Personen einen steigenden Anteil der Bevölkerung ausmacht.

In dieser Bachelorarbeit wurden Daten aus dem System „BicikeLJ“ der slowenischen Hauptstadt Ljubljana analysiert und interpretiert. Durch die elektronische Abwicklung der Ausleih- und Rückgabevorgänge kommen im Laufe eines Jahres große Datenmengen zusammen, sodass daraus fundierte Aussagen über das Verhalten der Benutzer getätigt werden können.

Die Dokumente mit den Datensätzen wurden im Rahmen der Bachelorarbeit vom Institut für Verkehrswissenschaften an der Technischen Universität Wien, Forschungsbereich für Verkehrsplanung und Verkehrstechnik, zur Verfügung gestellt.

1.1 Bike-Sharing-System BicikeLJ

Das Verleihsystem wurde im Mai 2011 in Betrieb genommen und bietet an 36 Stationen die Möglichkeit ein Fahrrad auszuborgen. Pro Terminal stehen 20 Leihfahräder zur Verfügung. Die Anzahl der Verleihstationen variierte in der bisherigen Geschichte von BicikeLJ. Darauf wird in der späteren Datenaufbereitung genauer eingegangen.

An den Terminals der Verleihstationen wird über ein im Vorhinein angelegtes Benutzerkonto ein Fahrrad nach dem Selfservice-Prinzip ausgeborgt. Die kürzest mögliche Ausleihzeit beträgt 5 Minuten, die maximale Ausleihdauer 24 Stunden. Dabei gelten folgende Tarife:

- Erste Stunde: kostenlos
- Zweite Stunde: 1,00 €
- Dritte Stunde: 2,00 €
- Jede weitere Stunde: 4,00 €

Falls der gewünschte Rückgabeort keinen freien Stellplatz hat, kann man sich am Terminal zusätzliche 15 Minuten kostenlos freischalten, um einen alternativen Standort zur Beendigung der Fahrt zu suchen.

Sollte es Probleme bei der Rückgabe eines Leihfahrrades geben wird empfohlen, es nach einer kurzen Wartezeit von 2 Minuten erneut zu probieren. Wenn es nach wie vor zu Komplikationen kommt, gibt es die Möglichkeit im Call Center anzurufen. [1][2]

1.2 Datenerfassung

Alle im System erfolgreich durchgeführten Ausleihvorgänge werden an den Terminals elektronisch verarbeitet und in Datenbanken abgespeichert. Die daraus resultierenden Excel-Dokumente umfassen immer den Zeitraum eines Kalendermonats.

Der Untersuchungszeitraum für diese Bachelorarbeit beträgt ungefähr ein Jahr. Genau betrachtet stammen die ersten Datensätze vom 01.08.2014 und die letzten vom 28.08.2015. Dadurch ergibt sich eine Betrachtungsdauer von ca. 13 Monaten bzw. 392 Tagen. In den Listen sind für jede Ausleihung folgende Informationen angeführt:

- Ausleihstation
- Datum und Uhrzeit der Ausleihung
- Nummer des Fahrrades
- Datum und Uhrzeit der Rückgabe
- Rückgabestation

1.3 Ziel der Analyse

Die Kernaufgabe bestand darin, die Datensätze aus dem Bike-Sharing-System BicikeLJ durch richtige Aufbereitung analysieren zu können. Die Vielzahl an Daten musste umgerechnet, geordnet und strukturiert aufgelistet bzw. gefiltert werden, um die gewünschten Informationen darstellen zu können. Anschließend sollten Aussagen über das Nutzerverhalten anhand von Kennzahlen quantifiziert werden und durch Interpretation wichtige Erkenntnisse hervorgehoben werden.

Zu den betrachteten Kennzahlen gehören Informationen wie Weglänge zwischen den Stationen (siehe Kapitel 3.1), Ausleihdauer (siehe Kapitel 3.2), durchschnittliche Reisegeschwindigkeit

(siehe Kapitel 3.3), häufig gefahrene Strecken (siehe Kapitel 3.4), Ausleih- und Rückgabehäufigkeiten (siehe Kapitel 3.5) an den verschiedenen Stationen und Aussagen über die Bilanz der einzelnen Stationen (siehe Kapitel 3.6).

Aufgrund der zur Verfügung gestellten Daten waren die Möglichkeiten der zu untersuchenden Zusammenhänge jedoch begrenzt. Zum Beispiel konnten keine Korrelationen berücksichtigt werden, die Nutzerinformationen wie Alter oder Geschlecht benötigen. Die Nummern der entlehnten Fahrräder wurden in den weiteren Fragestellungen ebenfalls nicht berücksichtigt.

Schließlich mussten sinnvolle Darstellungsmethoden gewählt werden, um die gewünschte Information einwandfrei und unverwechselbar ablesen zu können. Es wurden daher verschiedene Typen von Diagrammen und Karten mit den Programmen Microsoft Excel und QGIS gewählt.

2 Datenanalyse

2.1 Aufbereitung

Zu Beginn musste die Vielzahl an Dokumenten in einer Datenliste zusammengefasst werden, um die Gesamtheit der Datensätze betrachten zu können.

Einige Monate waren in zwei Dokumente getrennt. Bei den betroffenen Monaten wurde festgestellt, dass am 15. Kalendertag, an dem die Trennung erfolgte, Datensätze fehlten. Die Ursache konnte nicht eruiert werden.

In einer gesonderten Liste, welche mit den Monatslisten übermittelt wurde, war die Soll-Anzahl der Datensätze pro Monat aufgelistet. Verglichen mit den vorhandenen Datensätzen über den gesamten Betrachtungszeitraum waren weniger vorhanden als es sein sollten (**Tab. 1**). Diese Lücken resultierten nicht nur aus dem zuvor beschriebenen Fehler mit den getrennten Monaten. Woher dieser weitere Datenfehler kam, konnte ebenfalls nicht eruiert werden.

Auf eine Interpolation von Zwischenwerten wurde verzichtet, da alle Monate ungleichmäßig davon betroffen waren. Daher wurden nur die vorhandenen Datensätze analysiert und hiermit auf vorhandene Datenlücken hingewiesen. Gemessen an der Gesamtanzahl von Ausleihvorgängen mit mehr als 788.000 Datensätzen, entsprechen die rund 38.000 fehlenden Datensätze zirka 5 %. Die Auswirkungen von den Datenlücken (**Tab. 1**) werden mangels Alternativen jedoch vernachlässigt, da es keine Möglichkeit gibt, diese zu rekonstruieren.

Jan	Feb	Mär	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez	Σ		
08/2014 bis 12/2014							61 200	67 110	93 916	65 128	44 359	331 713	SOLL	
							59 821	63 213	89 177	64 560	43 953	320 724	IST	
							-1 379	-3 897	-4 739	-568	-406	-10 989	Δ	
37 864	29 404	60 437	70 202	83 306	83 206	68 388	62 059	01/2015 bis 08/2015				494 866	SOLL	
37 537	28 623	57 553	64 712	80 633	76 553	63 077	59 337					468 025	IST	
-327	-781	-2 884	-5 490	-2 673	-6 653	-5 311	-2 722					-26 841	Δ	
Σ							SOLL	826 579	100%					
							IST	788 749	95%					
							Δ	-37 830	5%					

Tab. 1: Vergleich: Anzahl der Datensätze

2.2 Strukturierung

Zur Vereinfachung der Beschriftung in den Diagrammen und Karten wurden den Standorten Nummern zugewiesen. Somit ergeben sich 38 Ausleihstationen. In der **Tab. 2** sind die Standorte mit ihren geografischen Koordinaten ersichtlich.

Wie sich im Laufe der Analyse herausstellte, sind im Betrachtungszeitraum für die Stationen mit den Nummer 37 und 38 keine Daten vorhanden. Es konnte nicht in Erfahrung gebracht werden, ob die Stationen bereits vor August 2014 aus dem Netz entfernt wurden oder ob sie erst nach August 2015 in das Stationsnetz aufgenommen wurden.

Die in der **Tab. 2** ersichtlichen Koordinaten der Terminals sind in Dezimalform angegeben. Die Entfernungen zwischen den Ausleihstationen wurden in weiterer Folge mit einer für dieses Koordinatenformat passenden mathematischen Formel berechnet (siehe Formel 2).

Im Programm QGIS musste lediglich das richtige Koordinatenformat WG84 eingestellt werden, damit die Punkte am richtigen Ort angezeigt werden. Dies war für die Erstellung von Karten erforderlich, um die Ergebnisse der Auswertung auch anhand ihres geografischen Auftretens zeigen zu können. Insbesondere für die Häufigkeit der gefahrenen Wege und die Häufigkeit der Benutzung der einzelnen Terminals war die Kartendarstellung anschaulicher als ein Diagramm.

Nr.	Name Standort	Latitude	Longitude	Nr.	Name Standort	Latitude	Longitude
1	PREŠERNOV TRG-PETKOVŠKOVO NABREŽJE	46,051367	14,506542	20	STARA CERKEV	46,063420	14,494588
2	POGACARJEV TRG-TRŽNICA	46,051093	14,507186	21	KINO ŠIŠKA	46,069280	14,489710
3	KONGRESNI TRG-ŠUBICEVA ULICA	46,050388	14,504623	22	ŠPICA	46,040213	14,512044
4	CANKARJEVA UL.-NAMA	46,052431	14,503257	23	BARJANSKA C.-CENTER STAREJŠIH TRNOVC	46,040810	14,499510
5	BREG	46,046498	14,505148	24	ZALOŠKA C.-GRABLOVICEVA C.	46,054410	14,529780
6	GRUDNOVO NABREŽJE-KARLOVŠKA C.	46,045748	14,506196	25	TRŽNICA MOSTE	46,055801	14,534156
7	MIKLOŠICEV PARK	46,054168	14,507060	26	ROŽNA DOLINA-ŠKRABCEVA UL.	46,051439	14,492730
8	BAVARSKI DVOR	46,056820	14,505510	27	DUNAJSKA C.-PS PETROL	46,065136	14,509112
9	TRG OF-KOLODVORSKA UL.	46,057421	14,510265	28	PLECNIKOV STADION	46,069420	14,510520
10	MASARYKOVA DDC	46,057630	14,514264	29	DUNAJSKA C.-PS MERCATOR	46,074193	14,511134
11	VILHARJEVA CESTA	46,060050	14,513020	30	VOJKOVA C.-BOŽICEVA C.	46,075860	14,519590
12	PARK NAVJE-ŽELEZNA CESTA	46,063610	14,513150	31	ŠPORTNI CENTER STOŽICE	46,082590	14,520160
13	TRG MDB	46,047565	14,495687	32	KOPRSKA ULICA	46,033408	14,482468
14	PARKIRIŠČE NUK 2-FF	46,046377	14,501256	33	MERCATOR CENTER ŠIŠKA	46,087079	14,475439
15	AMBROŽEV TRG	46,049877	14,516308	34	CITYPARK	46,068805	14,546257
16	GH ŠENTPETER-NJEGOŠEVA C.	46,052670	14,519280	35	BTC CITY - DVORANA A	46,065297	14,543996
17	ILIRSKA ULICA	46,053144	14,513479	36	BTC CITY - ATLANTIS	46,063081	14,547851
18	TRŽAŠKA C.-ILIRIJA	46,044629	14,486699	37	TRNOVO	46,027453	14,499540
19	TIVOLI	46,059520	14,498200	38	P + R BARJE	46,037872	14,500072

Tab. 2: Koordinaten der Terminals

2.3 Berechnungsmethoden

2.3.1 Anzahl Entfernungen

Die Anzahl aller Stationen ist 38. Da ein Weg in beide Richtungen fahrbar ist und dieselbe Entfernung aufweist, sowie der Weg zwischen ein- und derselben Ausleih- und Rückgabestelle gleich 0 ist, kommt es zu 703 möglichen Distanzen zwischen den Terminals. Die Ermittlung der Anzahl aller Relationen wird nun in Formel (1) vorgerechnet.

$$\begin{aligned}
 \text{Anzahl Entfernungen} &= \frac{\text{Anzahl Stationen}^2 - \text{Anzahl Wege zur selben Station}}{2} = \\
 &= \frac{(38 * 38) - 38}{2} = 703 \text{ mögliche Entfernungen zwischen Ausleihung und Rückgabe}
 \end{aligned}
 \tag{1}$$

Zur Berechnung der Distanz wurden die Koordinaten der Terminals herangezogen und mit der Haversine-Formel (2) ausgewertet (siehe Referenz [3]). Diese Formel ermöglicht die Ermittlung der Entfernung zweier Punkte unter Berücksichtigung des Erdradius, da die Erde genähert einer Kugeloberfläche entspricht. Die Formel lautet wie folgt:

$$d = 2 * R * \arcsin \left(\sqrt{\sin^2 \left(\frac{B_{lat} - A_{lat}}{2} \right) + \cos(A_{lat}) * \cos(B_{lat}) * \sin^2 \left(\frac{B_{lon} - A_{lon}}{2} \right)} \right) \quad (2)$$

d = Distanz Punkt A zu Punkt B [m]

R = Erdradius = 6371 km

(A_{lat} / A_{lon}) = Punkt A Koordinaten (Latitude / Longitude)

(B_{lat} / B_{lon}) = Punkt B Koordinaten (Latitude / Longitude)

2.3.2 Geschwindigkeit

Die Geschwindigkeiten zwischen den Entnahme- und Rückgabeorten ergeben sich aus den bekannten Verleihzeiten und den ermittelten Luftdistanzen. Dazu wurde die bekannte Beziehung von zurückgelegtem Weg in der dabei vergangenen Zeit aus der Kinematik herangezogen. [4] Es handelt sich bei diesen Zahlenwerten jedoch nicht um Durchschnittsgeschwindigkeiten, sondern um Reisegeschwindigkeiten vom Punkt der Ausleiherung zum Punkt der Rückgabe.

2.3.3 Dichtefunktion

Die Gauß'sche Glockenkurve wurde im Programm Microsoft Excel mit der Funktion der Normalverteilung berechnet. Für die Darstellung dieser Häufigkeitsverteilung ist es notwendig, die Standardabweichung, den Mittelwert und den Maximalwert der Kurve zu kennen. Nachdem diese Werte ebenfalls mit Excel-Funktionen berechnet wurden, konnten mit der Funktion der Normalverteilung Faktoren ermittelt werden. Multipliziert man diese Faktoren mit den Werten der Häufigkeitsverteilung in den jeweiligen Intervallen, erhält man die dazugehörigen Werte der Gauß'schen Glockenkurve. [5]

2.3.4 Bestimmtheitsmaß

Das Bestimmtheitsmaß R^2 stellt eine Kennzahl zur formalen Beurteilung der Anpassungsgüte einer Regression in der Statistik dar. Der Wert kann jede reelle Zahl zwischen 0 und 1 annehmen. Der Zahlenwert 1 bedeutet, dass eine unabhängige Variable anhand der Varianz sehr gut durch eine abhängige Variable vorhergesagt werden kann. Je weiter sich dieser Determinationskoeffizient 0 annähert, desto schlechter wird die Vorhersage einer Zufallsvariable. [6]

3 Ergebnisse der Analyse

3.1 Weglängen

Die Hauptstadt Sloweniens zählt mit einer Fläche von rund 275 km² und einer Einwohnerzahl von 288.250 (Stand 1. Juli 2017) zu den kleineren Hauptstädten Europas [7]. Die Innenstadt bildet den geografischen Mittelpunkt des BicikeLJ Netzes; nur einige wenige Stationen befinden sich am Stadtrand. Diese Aussage ist nicht in Zahlen quantifizierbar. Sie wird jedoch anhand der Grafiken im Anhang (siehe **Anhang 1** bzw. **Anhang 2**) ersichtlich.

Die Entfernungen zwischen allen Stationen zueinander wurden über Luftlinien ermittelt. Dies kommt daher, dass der tatsächlich gefahrene Weg eines betrachteten Datensatzes nicht bekannt ist. Kunden können mit ihren Leihfahrrädern jeden erdenklichen Weg von der Ausleihstation zur Rückgabestelle gefahren sein. Die berechneten Weglängen (Vorgangsweise siehe 2.3.1) wurden anschließend in Intervalle von jeweils einem Kilometer eingeteilt. Die auf zwei Nachkommastellen genauen Distanzen zwischen den Stationen sind lediglich für die Berechnung der durch-

schnittlichen Reisegeschwindigkeit notwendig. Eine grafische Darstellung in einem Diagramm (**Abb. 3**) erscheint nicht zweckmäßig.

Die am weitesten voneinander entfernten Stationen sind die Terminals 33 und 37. Ihre Entfernung beträgt rund 6.886 Meter. Die Verleihstationen 1 und 2 sind hingegen nur 58 Meter voneinander entfernt. Wie man aus diesen Beispielen erkennt, sind die Unterschiede der Distanzen markant. Dies wird im folgenden Balkendiagramm (**Abb. 1**) veranschaulicht.

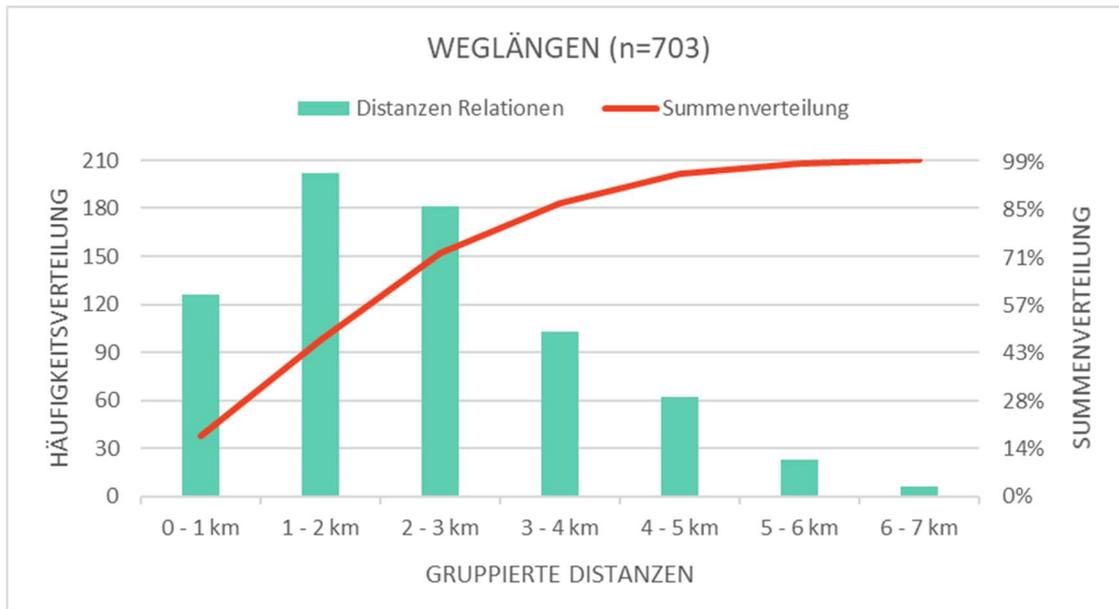


Abb. 1: Gliederung der Weglängen

Wie aus **Abb. 1** erkennbar, machen die Entfernungen von weniger als 4 km über 85 % aller Distanzen aus. Im Stadtzentrum sind die Ausleihstationen auf engem Raum situiert, dadurch ergibt sich der hohe Prozentsatz an kurzen Weglängen. Die am weitesten voneinander entfernten Terminals liegen jeweils auf den gegenüberliegenden Seiten der Innenstadt (siehe **Anhang 1** bzw. **Anhang 2**). Ihr Anteil beträgt nur wenige Prozent.

Um die Attraktivität des Bike-Sharing-Systems anhand der Entfernungen zwischen den Terminals bewerten zu können, müssten international mehrere Verleihsysteme miteinander verglichen werden. Ebenso wären Erhebungen über die Kundenmeinung erforderlich. Dies war nicht Teil der Fragestellung dieser Arbeit. Somit wurden die Weglängen in weiterer Folge nur für die Berechnung der Punkt-zu-Punkt-Geschwindigkeiten herangezogen.

3.2 Ausleihdauer

Da die Datensätze durch die elektronische Aufzeichnung exakte Zeitangaben aufweisen, konnte die Dauer der Ausleihvorgänge auf die Sekunde genau ermittelt werden. Diese hohe Genauigkeit ist bei der Datenauswertung nicht notwendig. Die Ausleihdauer wurde lediglich auf ganze Minuten gerundet. Die ermittelten Zeiten wurden in einem Balkendiagramm (**Abb. 2**) dargestellt.

Um im Diagramm eine Darstellung der Verteilungskurve zu erhalten, wurden die Zeitintervalle klein gehalten und mit 2 Minuten festgelegt. Größere Intervalle hätten zu einer verzerrten Sichtbarkeit der Ergebnisse geführt, da schon in den ersten drei Klassen unter 6 Minuten große Sprünge in der Häufigkeitsverteilung auftreten.

Auffallend waren die Verleihzeiten von mehr als einer Stunde, welche nur rund 2 % aller Ausleihdauern ausmachen und durch einen Sprung am Ende des Balkendiagramms deutlich sichtbar sind. Diese Verleihzeiten wurden in einem Überlaufintervall zusammengefasst. Die längste Ausleihdauer beträgt 23 Stunden 40 Minuten. Die im Bike-Sharing-System maximal zulässige Dauer für einen Verleihvorgang beträgt 24 Stunden.

Da die größte Distanz zwischen zwei Terminals rund 7 Kilometer beträgt (siehe 3.1 „Weglängen“), kann darauf geschlossen werden, dass Kunden während ihrer Fahrt einen oder mehrere Zwischenstopps eingelegt haben könnten. Dies wird anhand der anschließend erläuterten Geschwindigkeiten sichtbar (siehe 3.3 „Durchschnittliche Reisegeschwindigkeit“).

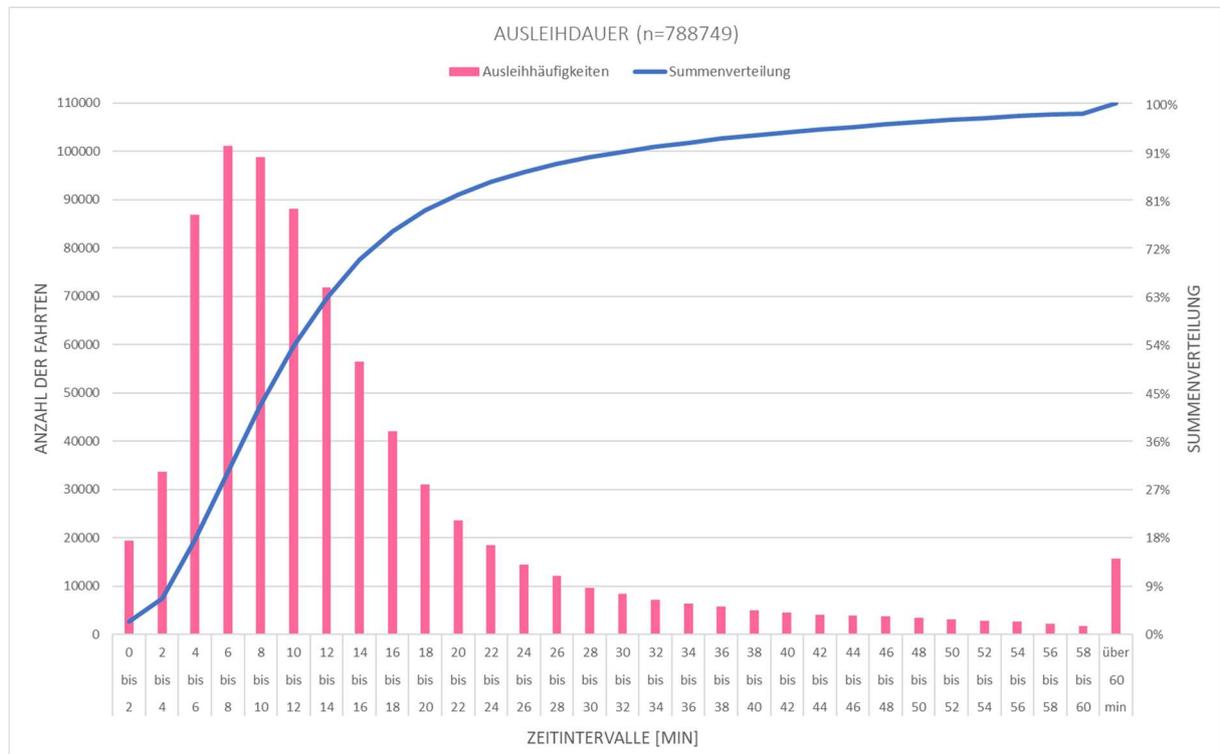


Abb. 2: Ausleihzeiten der Leihfahrräder

Aus **Abb. 2** ist deutlich erkennbar, dass 98 % aller Fahrten innerhalb einer Stunde wieder beendet werden. Das kann einerseits an den kurzen Distanzen im Netz liegen, andererseits aber auch am Preiskonzept (siehe 1.1 „Bike-Sharing-System BicikelJ“). Da die erste Stunde kostenlos ist, kann daraus geschlossen werden, dass viele Kunden die Fahrräder für schnelle Wege benutzen. Dies würde genau dem Gedanken eines Bike-Sharing-Systems entsprechen. Die im Kapitel 1.1 angeführte Preisliste für die verschiedenen Verleihzeiten könnte als Basis für die Berechnung der Einnahmen herangezogen werden. Wie zuvor erwähnt, dauern jedoch nur 2 % der Fahrten so lange, dass sie kostenpflichtig werden. Es wurde davon Abstand genommen, in dieser Richtung weitere Berechnungen anzustellen, da Rückschlüsse auf die Deckung der Kosten und damit die Erhaltung des Systems nicht möglich sind.

3.3 Durchschnittliche Reisegeschwindigkeit

Mittels der Distanzen zwischen den Terminals (Berechnung siehe 3.1 „Weglängen“) und der Ausleihzeiten (siehe Kapitel 3.2) wurden die vermutlichen durchschnittlichen Fahrgeschwindigkeiten ermittelt (Berechnung siehe 2.3.2 „Geschwindigkeit“).

Zur weiteren Interpretation wurden diese in einem Balkendiagramm mit blauer Farbe dargestellt. Die Verteilung der Geschwindigkeiten aus den Datensätzen sollte sich annähernd wie eine Gauß'sche Glockenkurve verhalten (siehe 2.3.3 „Dichtefunktion“). Diese Dichtefunktion wurde für die Anzahl von 788.749 Datensätzen, mit einem Mittelwert $\mu = 9,0$ km/h und einer Standardabweichung $\sigma = 2,24$ errechnet und mit einer orangefarbenen Linie im Diagramm eingezeichnet. Die graue Kurve stellt die Aufsummierung der blauen Balken in **Abb. 3**, also die Funktion der Summenverteilung, dar.

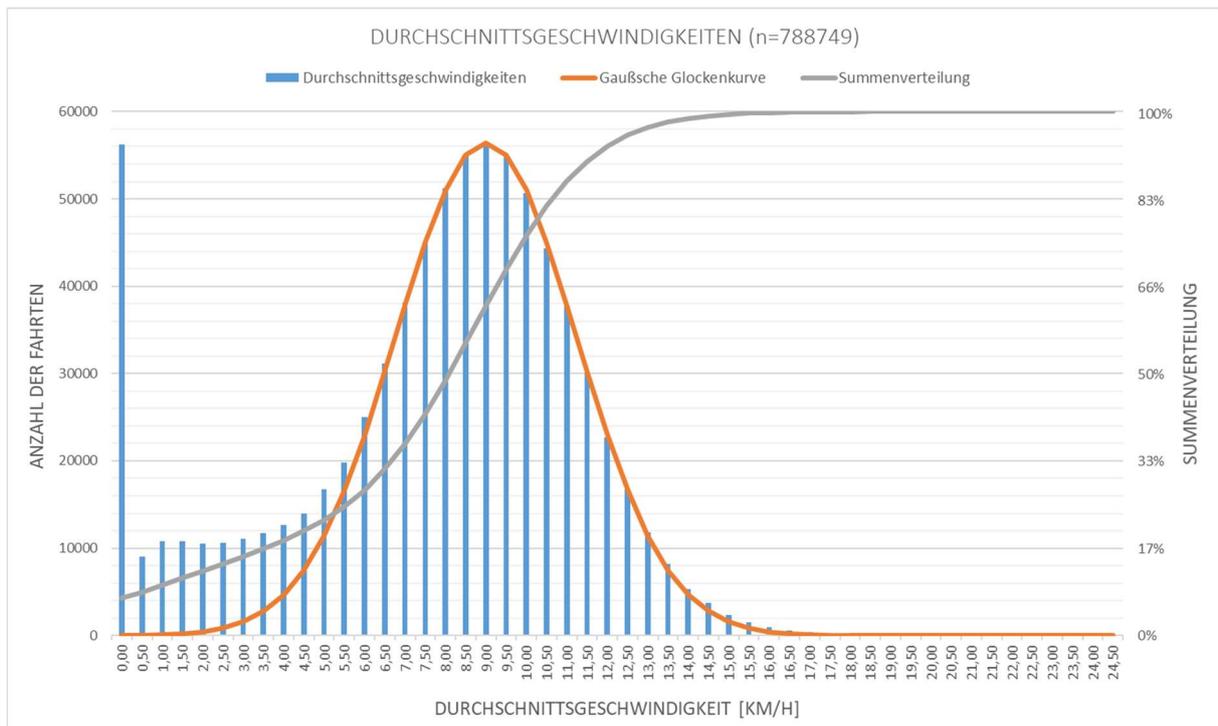


Abb. 3: durchschnittliche Punkt-zu-Punkt-Reisegeschwindigkeiten der Nutzer

Wie in **Abb. 3** ersichtlich, verhalten sich die blauen Balken der errechneten durchschnittlichen Punkt-zu-Punkt-Geschwindigkeiten nicht wie die orangefarbene Gauß'sche Glockenkurve. Auffallend sind die Werte der Intervalle unter 6 km/h, welche markant über der Dichtefunktion liegen. Ein Extremereignis stellt der Balken für die Fahrten mit einer rechnerischen Durchschnittsgeschwindigkeit von unter 0,5 km/h dar. Davon sind rund 7 % aller Fahrten betroffen. Eine Ursache hierfür könnte ein Zusammenhang zwischen vielen kurzen Wegen innerhalb des Stationsnetzes und den zugehörigen Ausleihzeiten sein. Möglicherweise ist ein Grund für die Verringerung der durchschnittlichen Reisegeschwindigkeiten auch die Nutzung der Leihfahrräder durch Touristen, die beim Sightseeing in der Stadt häufig stehen bleiben. In **Abb. 2** „Ausleihzeiten der Leihfahrräder“ ist anhand der blauen Kurve für die Summenverteilung erkennbar, dass exakt 80 % der Ausleihzeiten weniger als 20 Minuten dauern. Respektive dauern die verbleibenden 20 % zwischen 20 Minuten und 23 Stunden 40 Minuten. Daraus lässt sich auf Zwischenstopps während der Fahrten schließen, was wiederum zu einer Verringerung der Durchschnittsgeschwindigkeiten führt. Daher würde man bei einer weiteren Verwendung der Reisegeschwindigkeiten den Bereich der blauen Balken, welcher außerhalb der orangefarbenen Verteilungsdichtefunktion liegt, nicht berücksichtigen. Im Diagramm **Abb. 3** wäre somit der Geschwindigkeitsbereich unter 6 km/h unberücksichtigt.

3.4 Häufige Strecken

Die Datensätze wurden auch auf die Häufigkeit der gefahrenen Strecken ausgewertet. Die zugehörige Berechnung der Entfernungen wurde bereits unter 3.1 „Weglängen“ erklärt. Die Liste mit allen Datensätzen wurde auf die möglichen Distanzen zwischen Start- und Zielort gefiltert. Diese beinhalten mehr als die 703 Entfernungen (siehe 2.3.1 „Anzahl Entfernungen“), da hierbei Hin- und Rückweg getrennt betrachtet wurden. Wenn Ausleih- und Rückgabeort zusammenfallen, bedeutet das zweifelsohne nicht, dass ein Kunde nicht trotzdem eine große Distanz gefahren sein kann. Dadurch entstehen ebenfalls Fehler in der Auswertung, da der tatsächlich gefahrene Weg nicht bekannt ist. Aufgrund der großen Datenmenge haben diese jedoch keine große Relevanz. Es wurden daher die bekannten Luftdistanzen zwischen den 36 Ausleihstationen und die Häufigkeit der Benützung dieser Relationen genauer betrachtet.

Die zugehörige grafische Darstellung wurde mit QGIS erstellt. In Form eines Spinnennetzes sind die möglichen Wege zwischen den 36 Terminals ersichtlich. Es stellt alle Verbindungen im betrachteten Zeitraum von August 2014 bis August 2015 dar. Durch die Strichstärke der Linien und die angegebenen Zahlenwerte wird dargestellt, wie häufig eine Strecke pro Richtung und Jahr gefahren wurde. Die Farben und Skalierungen der Liniendicken sind in der Legende ersichtlich. Die Skalierung ist ein Vielfaches von 392 Tagen, das bedeutet z.B., dass in der Kategorie 0-392 ein Weg maximal ein Mal pro Tag oder seltener gefahren wurde. Ebenso sind die anderen Kategorien zu verstehen.

Es wurde davon Abstand genommen, die beiden richtungsabhängigen Häufigkeiten für einen Weg zu summieren. Die Häufigkeiten der beiden Fahrtrichtungen wurden in einer Kreuztabelle analysiert. Dabei kam man zu der Erkenntnis, dass es Unterschiede zwischen den beiden Richtungen eines Weges gibt. Diese sind interessanterweise in einem so geringen Ausmaß, dass sich der Linientyp aufgrund der Fahrten pro Tag nicht ändert (siehe zugehörige Legende in der Karte **Anhang 1**). Daher wurden die Häufigkeiten für beide Richtungen mit den gleichen Linientypen dargestellt. Die Grafik zeigt daher nur die Anzahl der Fahrten im Betrachtungszeitraum pro Richtung.

Auffällig ist das sternförmige Ergebnis der wichtigsten Wege. Daraus lässt sich schließen, dass viele Kunden ein Leihrad benutzen, um vom Stadtrand in die Innenstadt zu fahren oder umgekehrt. Die Grafik in **Anhang 1** veranschaulicht eine radiale Beziehungsstruktur. Am stärksten ist das anhand der hellblauen Linien sichtbar. Dieses grafische Ergebnis unterstützt die nicht eindeutig quantifizierbare Schlussfolgerung der häufig gefahrenen kurzen Wege in Stadtzentrum, welche in den vorherigen Kapiteln mehrmals betont wurde.

Deutlich ist, wie schon öfter erwähnt, dass die Terminals mit den Nummern 37 und 38 in den Datensätzen nicht vorkommen und daher keine Verbindungen zu anderen Stationen haben. Der Vollständigkeit halber wurden sie jedoch in die Grafik aufgenommen.

3.5 Ausleih- und Rückgabehäufigkeiten

Die Terminals des Bike-Sharing-Systems BicikeLJ sind unterschiedlich häufig frequentiert. Dies hat vermutlich mit der Lage, bezogen auf wichtige Verkehrsknotenpunkte in der Innenstadt, öffentliche Einrichtungen, sowie andere Plätzen mit hohem Menschenaufkommen zu tun.

Um einen Überblick zu bekommen, wie stark die Ausleih- und Rückgabehäufigkeiten der einzelnen Stationen untereinander variieren, wurden ein Diagramm mit Werten und eine Heatmap (siehe **Anhang 2**) mit farblicher Kennzeichnung erstellt. Die Differenzen wurden im Kapitel 3.6 „Bilanzen der Stationen“ erläutert.

Im Diagramm (**Abb. 4**) wird sichtbar, dass die Benützungsfrequenz der Stationen stark variiert. Die meisten Verleihereignisse weist der Terminal 1 mit 53.309 Ausleih- und 51.815 Rückgabevorgängen auf. Die wenigsten Fahrräder werden am Terminal 32 ausborgt. Hier finden lediglich 2.090 Ausleih- und 2.371 Rückgabevorgänge statt. Am Terminal 1 werden daher rund 25 Mal mehr Leihräder entlehnt und rund 22 Mal mehr Fahrräder abgestellt als beim Terminal 32.

Wie die Frequentierung mit der geografischen Lage in Zusammenhang steht, wird in einer Heatmap im Anhang dargestellt. Diese wurde mit QGIS auf Basis der ausgewerteten Datensätze erstellt. Die Legende darunter gibt eine Erklärung über die verwendeten Farben.

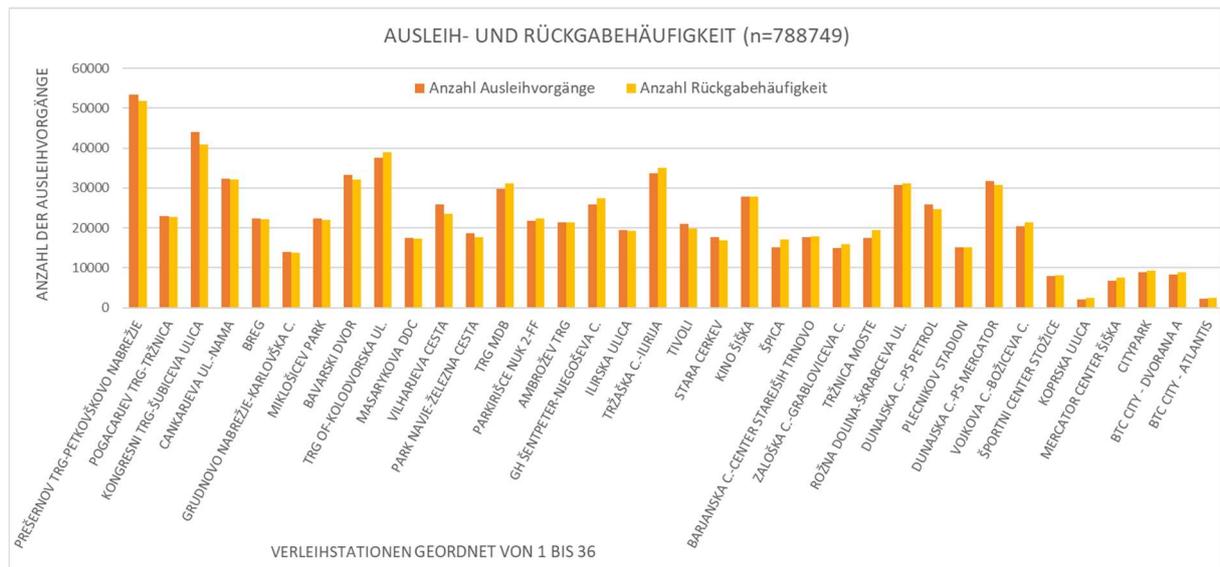


Abb. 4: Stationen und ihre Ausleihhäufigkeiten

Auf eine Auswertung der Rückgabehäufigkeit im **Anhang 2** „Heatmap der Ausleihhäufigkeiten“ wurde verzichtet, da bei der Betrachtung der Bilanzen (siehe Kapitel 3.6) der einzelnen Terminals ersichtlich wird, dass es nur geringfügige Abweichungen von der 1. Mediane gibt. Daher würde sich die Darstellung der Rückgabehäufigkeit mittels einer Heatmap nicht wesentlich von jener der Ausleihhäufigkeit unterscheiden.

Die Größe der Kreise in der Heatmap (siehe **Anhang 2**) hat keine Aussage. Der Radius wurde so festgelegt, dass eine gute Sichtbarkeit der Farben im Stadtzentrum garantiert werden kann.

3.6 Bilanzen der Stationen

Die Verleihstationen bei Bike-Sharing-Systemen weisen meist keine ausgeglichenen Bilanzen an Ausleih- und Rückgabevorgängen (siehe Kapitel 3.5) auf. Ob an einem Terminal mehr Fahrten starten oder enden, ist von einer Vielzahl an Faktoren abhängig, welche sich durch Bedeutung eines Standortes unterscheiden. Eine Quelle weist einen Überschuss an Ausleihvorgängen auf, eine Senke einen Überschuss an Rückgabevorgängen. Die Differenzen an Vorgängen müssen manuell durch den Betreiber ausgeglichen werden, da sonst die verschiedenen Terminals entweder dauerhaft überfüllt oder ständig leer sind.

Dieses Phänomen tritt vermehrt an Standorten auf, an denen große Menschenmengen unterwegs sind. Ein Beispiel wäre ein Bahnhof, von welchem die öffentliche Verkehrsanbindung zur Weiterfahrt in die Innenstadt unzureichend ist. Dies ist ein Extremereignis in den Morgenstunden aufgrund von Pendlern, welche den Terminal innerhalb eines kurzen Zeitraumes leeren. Jedoch kommen im Laufe des Tages nicht alle Leihräder wieder zurück an ihren Ausgangsort. So kommt es zu einem massiven Defizit an dem betroffenen Standort und ein Eingreifen des Betreibers ist erforderlich.

Anhand einer Gegenüberstellung der Ausleih- und Rückgabevorgänge kann man die Größenordnungen der positiven oder negativen Differenzen der Terminals zueinander schnell überblicken. Sie wurden in der folgenden Grafik (**Abb. 5**) dargestellt. Die Verleihstationen werden mit einem Punkt in einem Koordinatensystem markiert. Die dafür herangezogenen Koordinaten sind für die Ausleihvorgänge auf der y-Achse und für die Rückgabehäufigkeit auf der x-Achse.

Die graue Diagonale in der **Abb. 5** markiert die 1. Mediane, also eine ausgeglichene Bilanz zwischen Ausleih- und Rückgabevorgängen. Ihre Funktion lautet $y = x$. Oberhalb dieser befinden sich die Quellen, unterhalb die Senken. Die blaue Linie im Diagramm zeigt die Trendlinie aller Stationsbilanzen, welche sich verschieden zur grauen Diagonale verhält. Die Funktion der Trendlinie lautet $y = 1,0271x - 583,49$. Die größten und kleinsten Differenzen von der Trendlinie

sind in der Grafik gesondert angeführt. Der Terminal 36 gehört zu den am geringsten frequentierten und hat mit einem Defizit von 129 Verleihvorgängen eine sehr ausgeglichene Bilanz. Nur Station 28 hat mit +80 Verleihvorgängen eine noch bessere Ausgeglichenheit. Im Diagramm (**Abb. 5**) liegt Terminal 22 sehr nahe bei Nummer 28, weist hingegen mit -1.961 Verleihvorgängen die größte Differenz aller Senken auf. Die mit Abstand größten Unterschiede von der Trendlinie hat Terminal 3, welcher ein Plus von 2.972 Verleihvorgängen aufweist. Ebenfalls auffällig sind die Bilanzen der Verleihstationen in der Innenstadt. Die Terminals 1 und 3 liegen deutlich im Quellen-Bereich, die in nächster Nähe befindlichen Terminals mit den Nummern 2, 4 und 7 bilanzieren hingegen ausgeglichen.

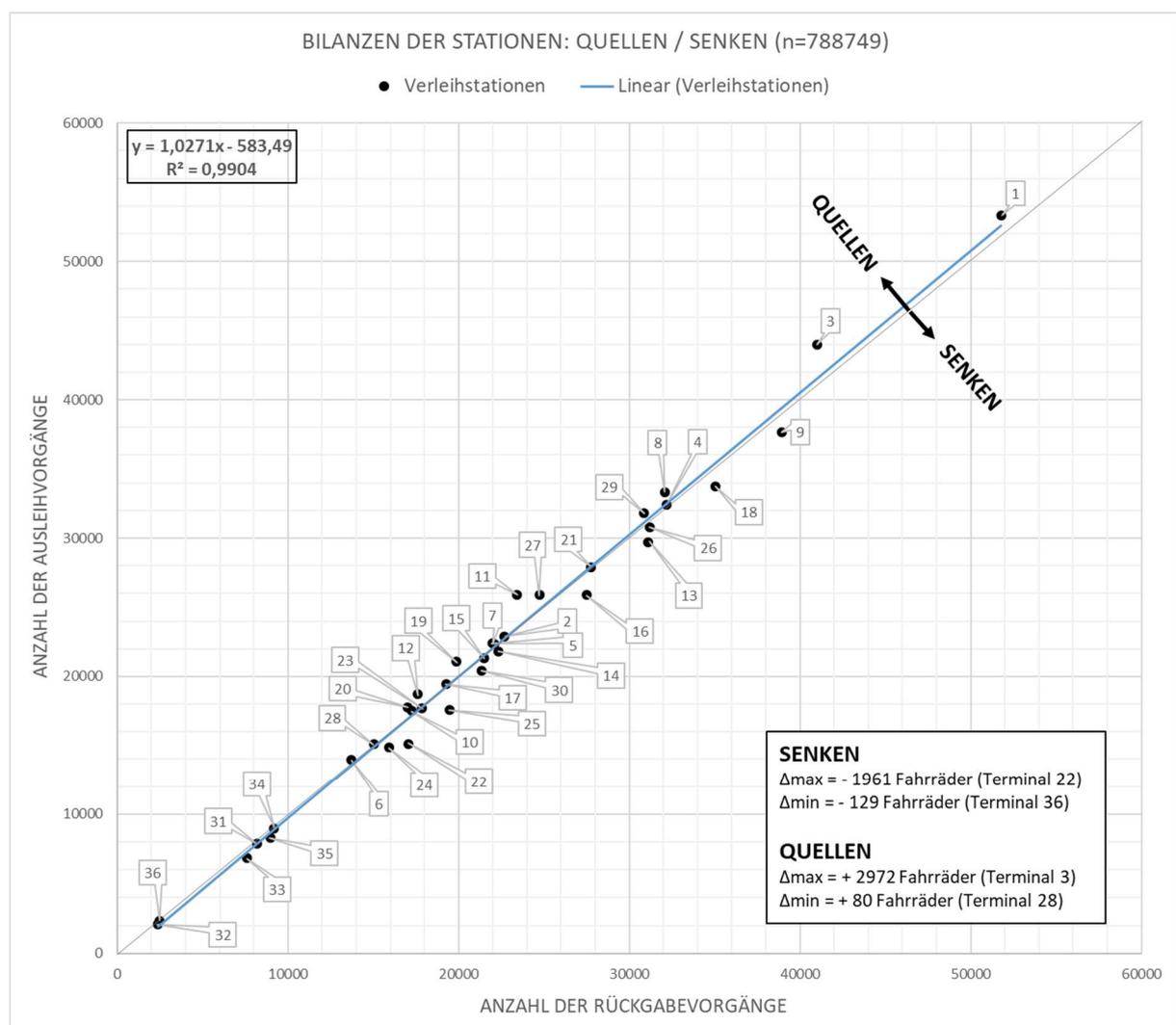


Abb. 5: Vergleich der Ausleih- und Rückgabebelastungen an den Terminals

In dieser Abbildung ist ersichtlich, wie groß die Unterschiede zwischen den Prozessen an den verschiedenen Terminals sind. Dies wird vor allem durch die Größenordnung der Achsen-skalierung sichtbar, deren Intervalle 10.000 Einheiten entsprechen. Das Bestimmtheitsmaß R^2 der Trendlinie im Diagramm hat einen Wert, der nahezu 1 entspricht. Das bedeutet, dass die Trendlinie sehr exakt die Punkte im Koordinatensystem abbildet (Erläuterung unter 2.3.4 „Bestimmtheitsmaß“).

Weitere Evaluierungen auf der Ebene der Auslastung der Verleihterminals konnten aufgrund mangelnder Informationen nicht durchgeführt werden.

4 Conclusio & Perspektive

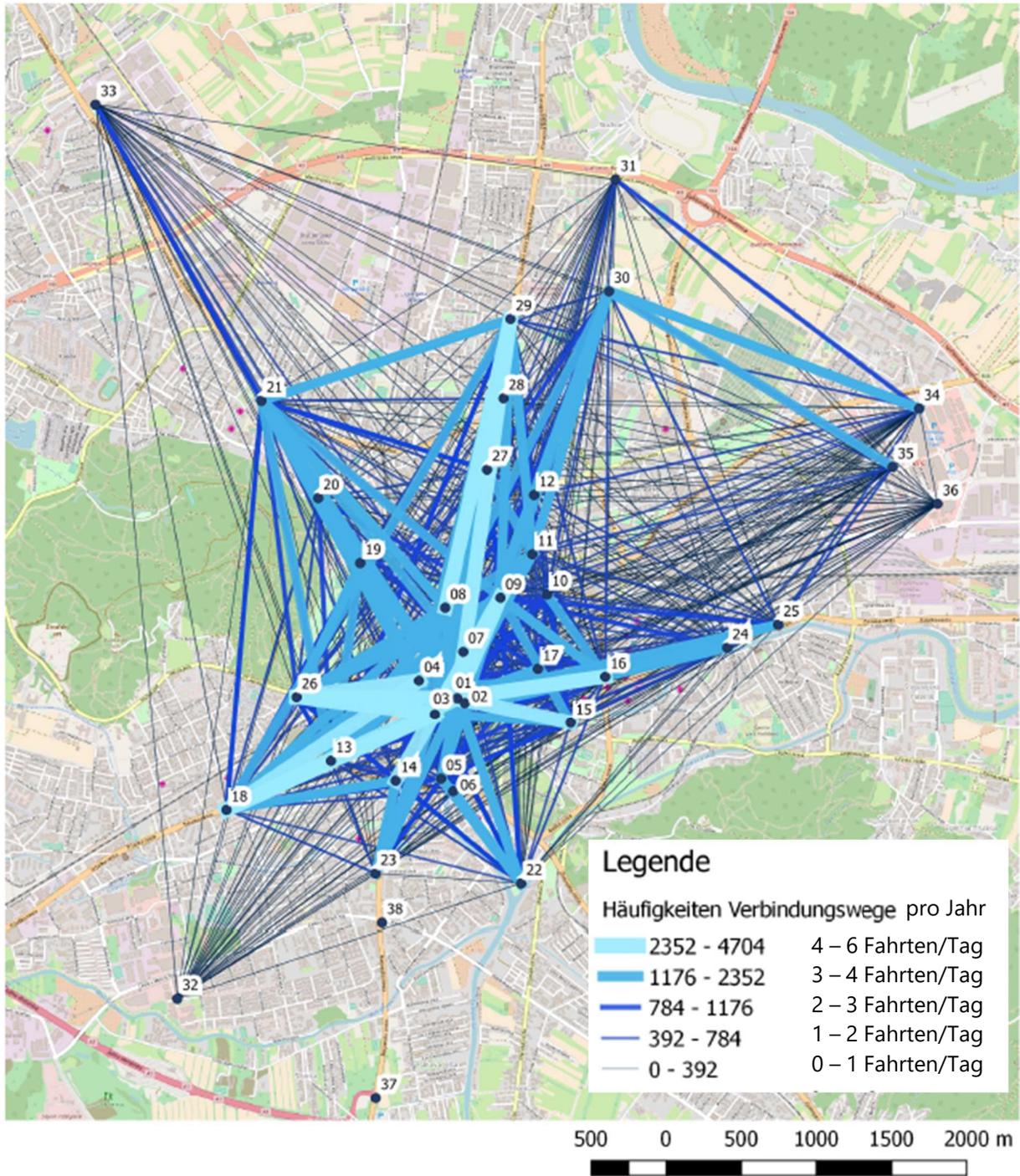
Die eingangs erklärten fehlenden Datensätze verzerren zweifellos die Diagramme und Karten. In welchem Ausmaß sich diese Datenlücken tatsächlich auswirken, ist allerdings unerheblich, da eine Rekonstruktion nicht möglich war und eine Interpolation nicht als zielführend erachtet wurde. Durch die große Anzahl an Datensätzen war es daher trotzdem möglich, sehr genaue Ergebnisse zu erhalten. Folglich konnten die Grafiken und Karten einwandfrei interpretiert und die zur Verfügung gestellten Informationen bestmöglich aufgearbeitet werden.

Die in der Arbeit gewonnenen Erkenntnisse über das Bike-Sharing-System BicikelJ bilden eine fundierte Basis für weiterführende Untersuchungen. Um exakte Aussagen über das Nutzerverhalten tätigen zu können, wäre es notwendig, Alter und Geschlecht der Benutzer zu kennen. Möglicherweise stößt solch eine Maßnahme aber an datenschutzrechtliche Grenzen. Andere Evaluierungen könnten die tatsächlich zurückgelegten Wegstrecken umfassen. Hierfür wäre eine Nachverfolgung der Leihfahräder mittels GPS erforderlich. Dadurch wäre es auch möglich, die errechneten Reisegeschwindigkeiten durch exakte Durchschnittsgeschwindigkeiten zu ersetzen. Mit diesen Werten würde auch eine neue Dichtefunktion die bisherige Näherung in **Abb. 3** „durchschnittliche Punkt-zu-Punkt-Reisegeschwindigkeiten der Nutzer“ ablösen.

Literaturverzeichnis

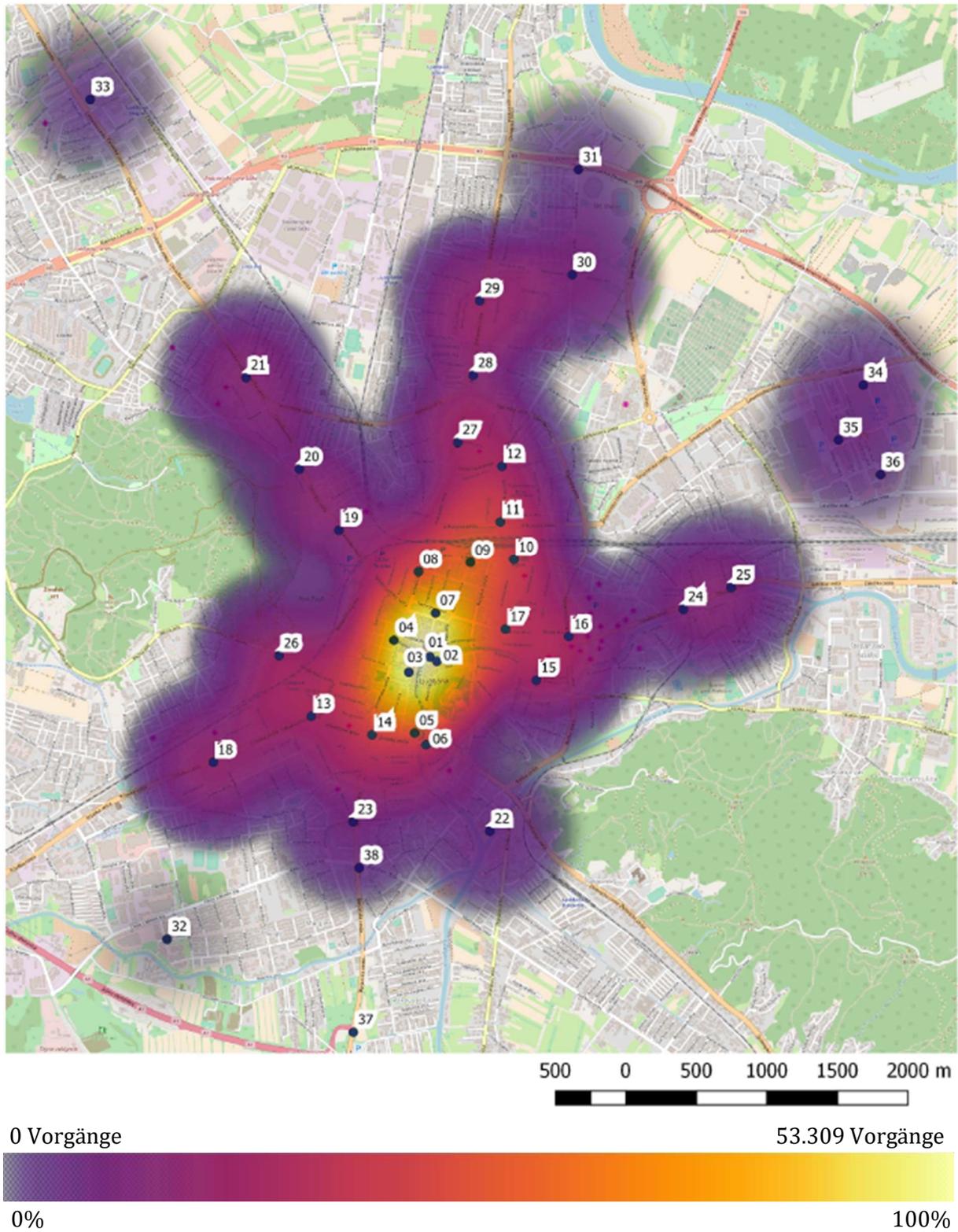
- [1] **Tourismus Homepage Ljubljana, Städtisches Fahrradverleihsystem BicikelJ**
<https://www.visitljubljana.com/de/besucher/ljubljana-und-region/verkehr-und-befoerderung/bicikelj/>
- [2] **Offizielle BicikelJ Homepage, Funktionsweise der Benützung**
<http://en.bicikelj.si/How-does-it-work>
- [3] **Haversine Formel zur Berechnung der Distanz zweier Punkte auf einer Kugeloberfläche**
https://en.wikipedia.org/wiki/Haversine_formula
- [4] **Geschwindigkeit, Kinematik**
<https://de.wikipedia.org/wiki/Geschwindigkeit>
- [5] **Glockenkurve, Dichtefunktion**
<https://de.wikipedia.org/wiki/Normalverteilung>
- [6] **Bestimmtheitsmaß, Interpretation der Kennzahl**
<https://de.wikipedia.org/wiki/Bestimmtheitsma%C3%9F>
- [7] **Ljubljana, Slowenien**
<https://de.wikipedia.org/wiki/Ljubljana>

Anhang 1:



Anhang 1: Netzkarte aller Relationen mit den Häufigkeiten pro Richtung pro Jahr (Luftlinie)

Anhang 2:



Anhang 2: Heatmap der Ausleihhäufigkeiten