



دانشگاه صنعتی شریف  
دانشکده مهندسی صنایع

عنوان:

## گزارش فاز دوم پروژه

اعضای گروه

۹۹۱۰۹۹۹۳ مهدی مهری  
۴۰۰۱۰۳۵۰۵ کیمیا ضیائی  
۴۰۰۱۰۳۶۰۲ محمد سبحان کسائی

نام درس

برنامه‌ریزی حمل و نقل

نام استاد درس

دکتر نائبی

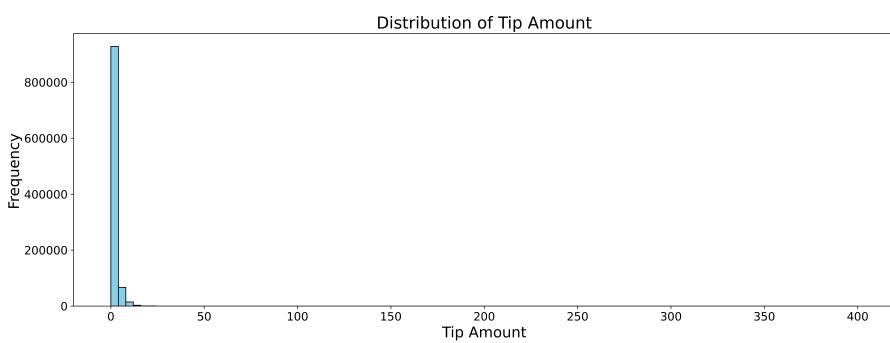
نیمسال اول ۱۴۰۳-۱۴۰۴

# ۱ مقدمات

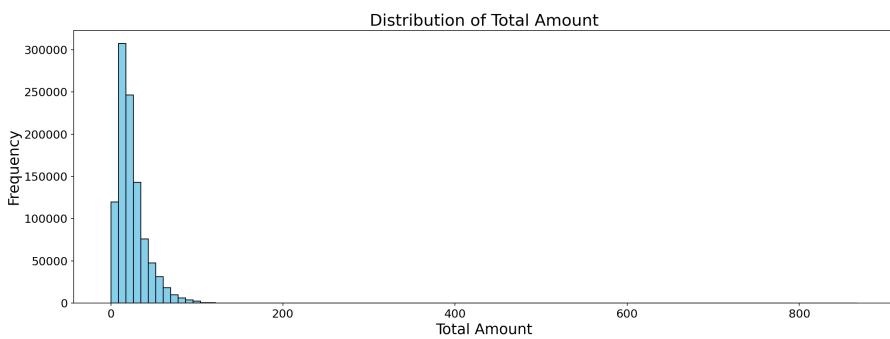
## ۱-۱ پاکسازی داده‌ها

در این قسمت، به مانند فاز اول، دیتاهایی که معتبر نیستند، مانند مقدار انعام یا هزینه سفر منفی، فاصله بسیار زیاد برای سفر، فاصله صفر برای سفر و غیره را حذف کرده و به جای آن مقادیر Null قرار دادیم. همچنین تنها داده‌های مربوط به سال ۲۰۲۱ را نگه می‌داریم.

## ۲-۱ مدیریت داده‌های ناموجود



شکل ۱: توزیع مقدار انعام سفر



شکل ۲: توزیع هزینه کل سفر

با توجه به اینکه مقادیر Null برای انعام و هزینه سفر کمتر از ۵ درصد است، باید مقادیر Null را با میانگین یا میانه آن جایگزین کنیم. همانطور که در نمودارهای بالا مشاهده می‌شود، با توجه به چلگی موجود در توزیع انعام و هزینه سفر، و دنباله طولانی که در توزیع این دو مشاهده می‌شود، مقادیر بسیار زیاد می‌تواند بر میانگین این دو متغیر تاثیر بگذارد، بنابراین بهتر است مقادیر Null را با میانه این دو متغیر جایگزین کنیم.

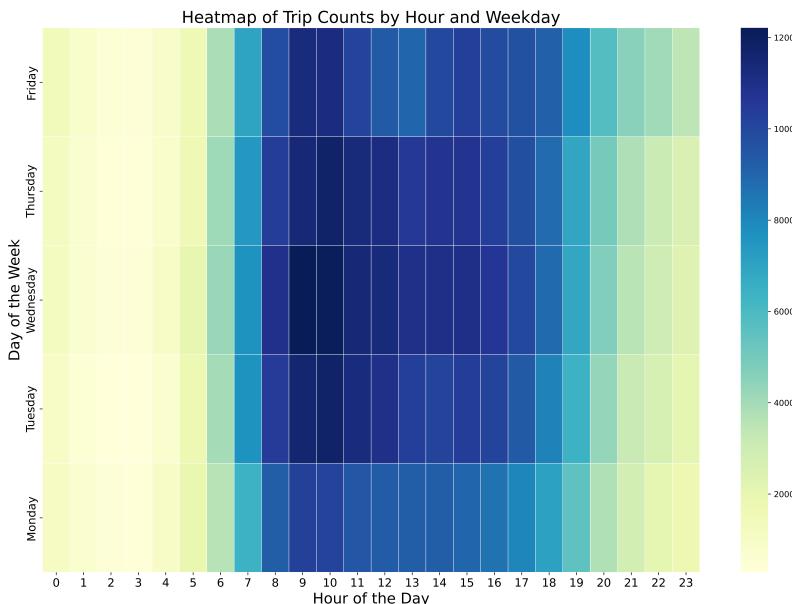
نیاز به پر کردن NULL برای باقی متغیرها نیست، زیرا در این فاز به مقادیر باقی متغیرها نیازی

نداریم.

### ۳-۱ پیش‌پردازش داده‌ها

در این قسمت، علاوه بر تغییر متغیرهای مربوطه مانند روزهای هفته، روزهای تعطیل و آخر هفته‌ها (شنبه و یکشنبه) را از داده‌ها حذف می‌کنیم، زیرا که در دام پروژه ذکر شده فقط روزهای کاری مد نظر قرار گیرد.

### ۴-۱ گروه‌بندی زمانی



شکل ۳: نقشه گرمایی تعداد سفرها در هر ساعت

در این قسمت، با رسم Heatmap تعداد سفرهای انجام شده در ساعات مختلف روز، هر روز را به سه گروه مختلف تقسیم می‌کنیم:

- ۱۲ شب تا ۶ صبح: کمترین مقدار سفرها در این زمان اتفاق می‌افتد.
- ۷ صبح تا ۷ شب: با توجه به ساعات کاری و فعالیت‌های روزانه، بیشترین مقدار سفرها در این بازه زمانی اتفاق می‌افتد.
- ۸ شب تا ۱۱ شب: تعداد سفرها در این بازه زمانی در مقایسه با دو بازه زمانی ذکر شده متوسط

است.

## ۲ خواسته اول

### ۱-۲ سوال اول

ماتریس OD خواسته شده ترسیم گردید و نتیجه به صورت زیر می باشد:



شکل ۴: ماتریس OD

### ۲-۲ سوال دوم

در این قسمت ابتدا مقادیر روی قطر اصلی را برابر صفر قرار می دهیم، سپس همه مقادیر ماتریس OD را مرتب می کنیم.

### ۳-۲ سوال سوم

همانطور که از ماتریس OD ترسیمی مشخص است، بیشترین سفرهای صورت گرفته درون یک بخش به ترتیب مربوط به بخش های زیر می شود:

Manhattan —

Brooklyn —

Queens —

Bronx —

Sataten Island —

EWR —

## ۴-۲ سوال چهارم

### ۱-۴-۲ تخصیص تاکسی‌ها به نواحی

همانطور که از ماتریس OD مشخص است، حجم قابل توجهی از سفرها به سفرهای درون ناحیه‌ای اختصاص دارد. و به ترتیب Manhattan، Brooklyn و Queens بیشترین سفرها را به خود اختصاص می‌دهد. بنابراین تعداد بیشتری از تاکسی‌ها باید در این سه بخش مستقر باشند، و بیشتر تاکسی‌ها باید مختص سفرهای درون ناحیه‌ای باشند. همچنین به طور کلی برای تخصیص سفرها به باقی نواحی نیز باید به ماتریس OD توجه داشت.

### ۲-۴-۲ تخصیص تاکسی‌ها از نظر زمانی

همانطور که از heatmap مشخص است، از ساعت ۰ تا ۵ بامداد تعداد سفرها بسیار کمتر است، بنابراین به تاکسی‌های کمتری در این ساعات نیاز داریم. از ساعت ۶ رفته تعداد سفرها افزایش می‌یابد، و تا ساعت ۸ تعداد سفرها به اوج خود می‌رسد. بنابراین تعداد تاکسی‌ها را نیز باید از ساعت ۶ به تدریج افزایش دهیم، و تا ساعت ۸ بیشترین ظرفیت از لحظه تعداد تاکسی‌ها در دسترس باشد. همچنین از ساعت ۱۹ می‌توان تعداد تاکسی‌ها را به تدریج کاهش داد، که تا ساعت ۲۳ ظرفیت ناوگان حمل و نقل به کمترین مقدار خود برسد. البته این موضوع نیز باید مد نظر قرار گیرد که در روزهای جمعه، از آنجایی که فردای این روز تعطیل است، حجم سفرها در ساعت ۱۹ و ۲۰ به تدریج کاهش می‌یابد، اما در ساعت ۲۱ تا ۲۳ حجم سفرها ثابت می‌ماند. بنابراین در روزهای جمعه نیز تعداد تاکسی‌ها از ساعت ۱۹ باید کاهش یابد، اما در ساعت پایانی شب همچنان باید تعداد تاکسی‌های بیشتری نسبت به باقی روزها در دسترس باشد.

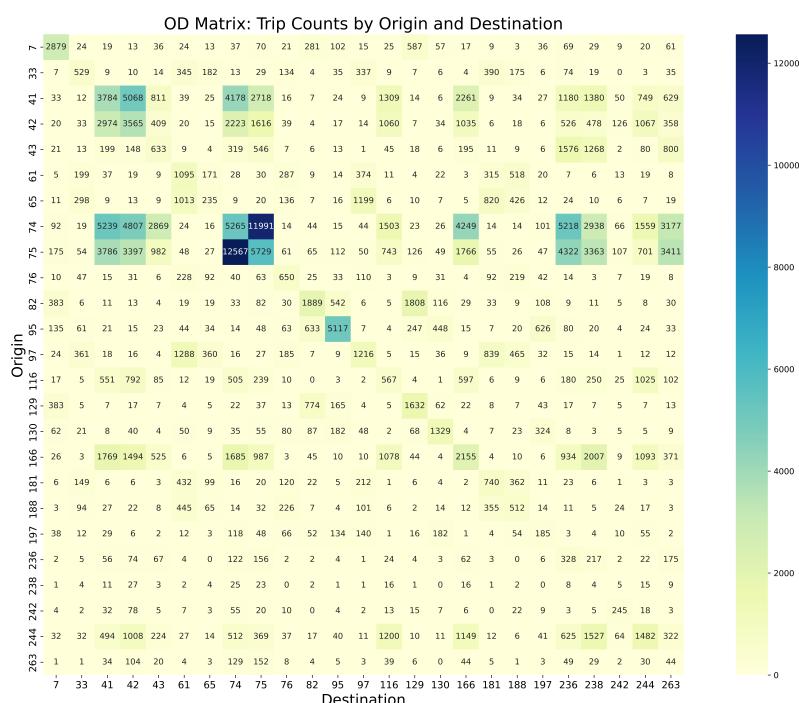
## ۳-۴-۲ قیمت‌گذاری

در صورتی که در ساعتی از روز تعداد تاکسی‌های موجود پاسخگوی نیاز مسافران نباشد، و یا با تقاضای زیاد و غیرقابل پیش‌بینی مواجه شدیم، قیمت‌گذاری به صورت پویا باید صورت گیرد، تا با افزایش قیمت راننده‌های بیشتری مایل به جابجایی مسافر باشند، و مسافرانی که واقعاً به تاکسی نیاز دارند از آن استفاده کنند، ممکن است با افزایش قیمت باقی افراد حمل و نقل عمومی را برای سفر خود انتخاب کنند.

## ۳ خواسته دوم

### ۱-۳ تشکیل ماتریس‌ها

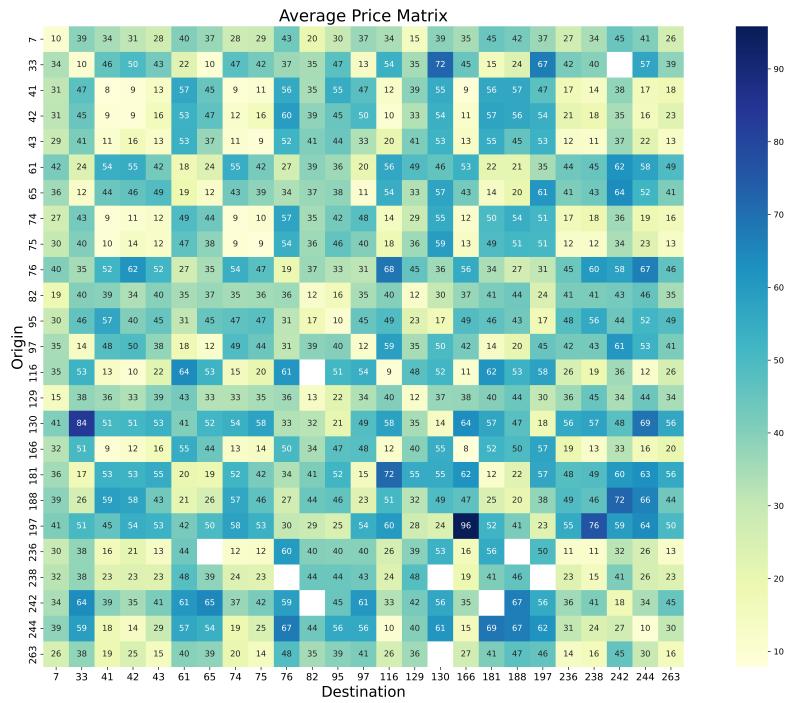
در این قسمت ابتدا ماتریس OD را برای همه ۲۶۵ منطقه رسم می‌کنیم، سپس ۲۵ منطقه که برای آن‌ها مجموع سفرهای تولید و جذب شده از باقی مناطق بیشتر است را انتخاب می‌کنیم و زیرماتریس OD را برای آن‌ها تشکیل می‌دهیم.



شکل ۵: زیرماتریس OD

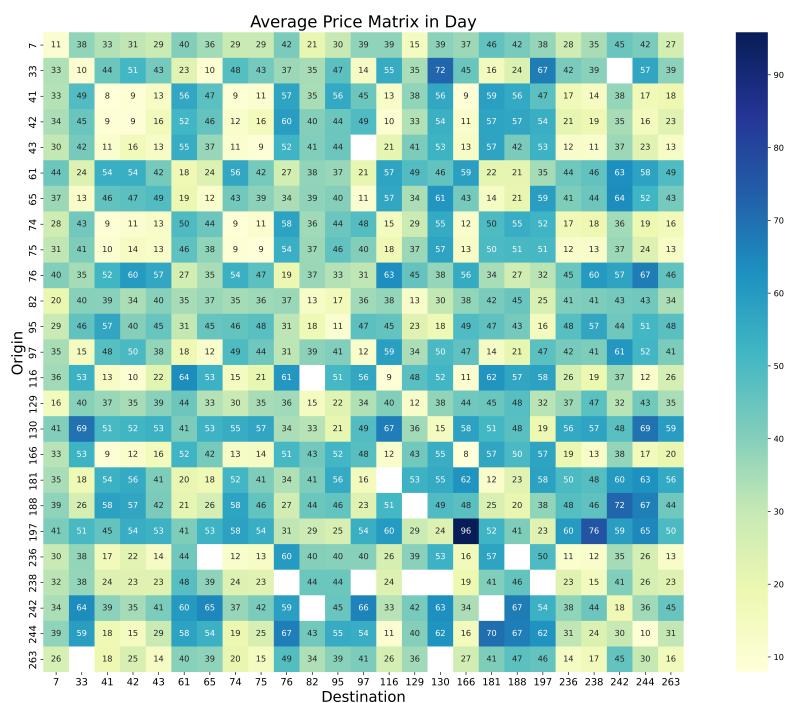
سپس ماتریس قیمت را برای ۲۵ منطقه مذکور، با میانگین گرفتن از متغیر Total Amount به ازای سفرهای صورت گرفته از هر منطقه به منطقه دیگر تشکیل می‌دهیم. این ماتریس را به

تفکیک هر یک از سه گروه زمانی که پیشتر ذکر کردیم نیز تشکیل می‌دهیم.



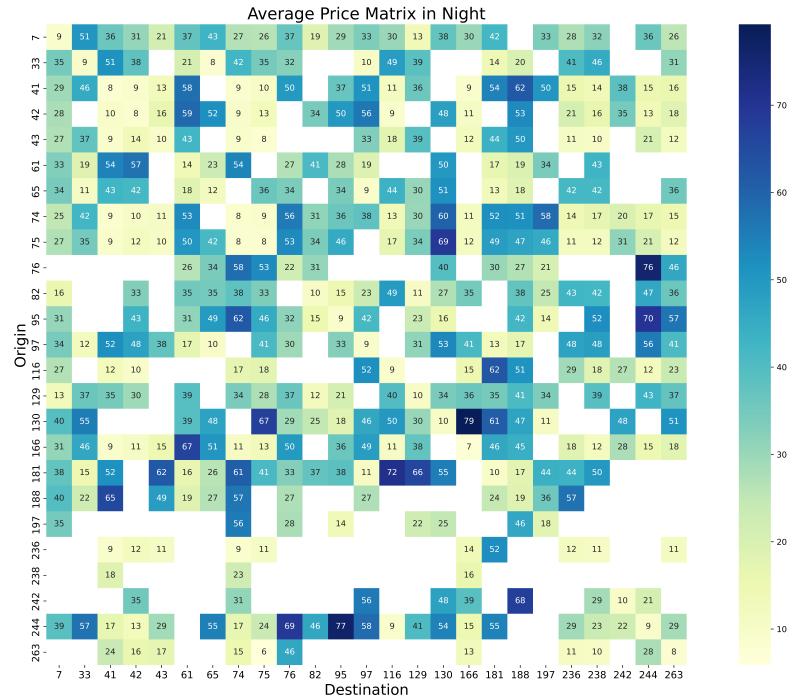
شکل ۶: زیرماتریس هزینه

زیرماتریس OD در بازه زمانی ۷ صبح تا ۷ شب به شکل زیر است:



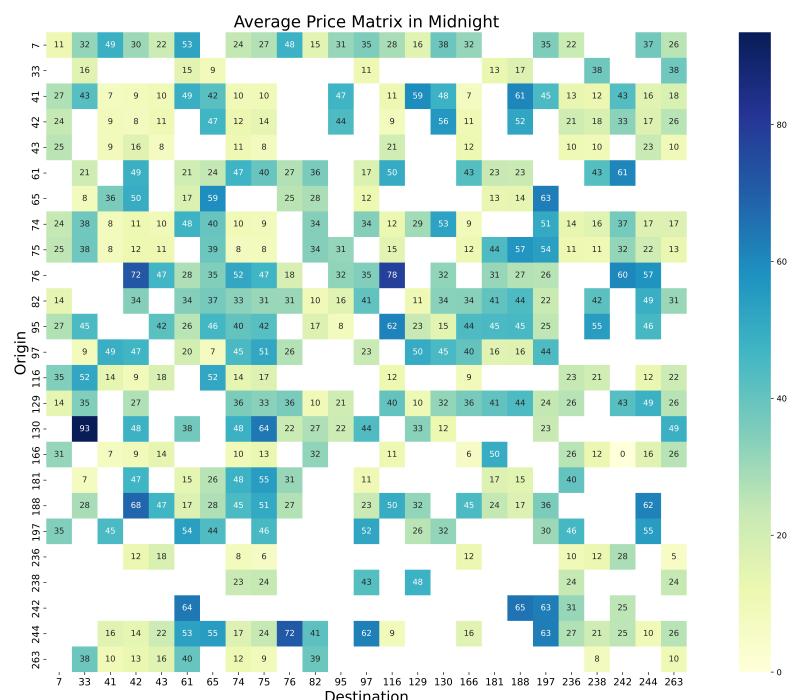
شکل 7: زیرماتریس هزینه در بازه زمانی روز

زیرماتریس OD در بازه زمانی ۸ شب تا ۱۱ شب به شکل زیر است:



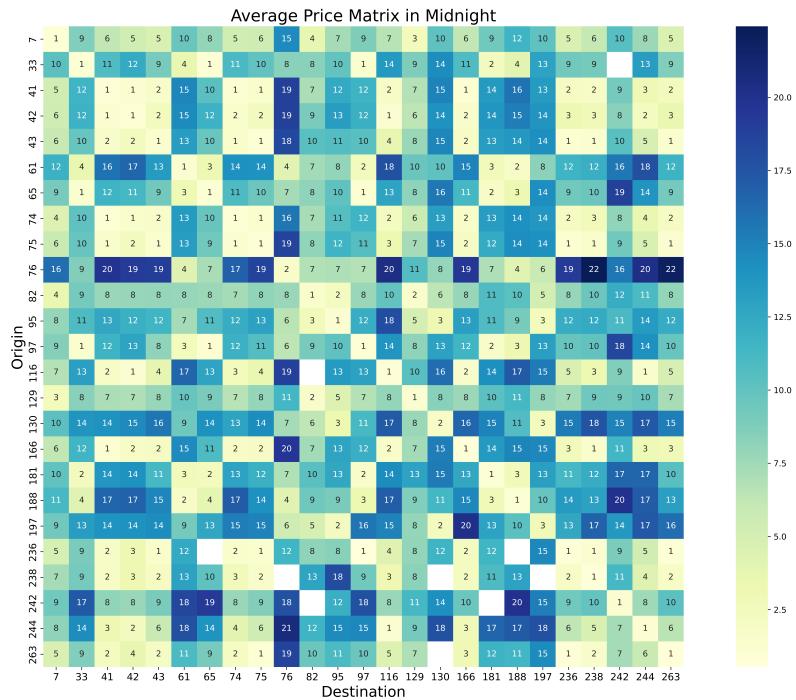
شکل ۸: زیرماتریس هزینه در بازه زمانی شب

زیرماتریس OD در بازه زمانی ۱۲ شب تا ۶ صبح به شکل زیر است:



شکل ۹: زیرماتریس هزینه در بازه زمانی نیمه شب

در نهایت ماتریس فاصله را به عنوان تخمینی از فاصله بین مناطق، با میانگین گرفتن بر طول سفرهای صورت گرفته از هر منطقه به منطقه دیگر به دست می‌آوریم.



شکل ۱۰: زیرماتریس فاصله

## ۲-۳ مدل‌سازی

در این قسمت، ابتدا به مدل‌سازی خواسته سوال پرداخته می‌شود و سپس در ادامه، به پاسخ سوالات موجود در فایل پروژه مربوط به خواسته دوم پرداخته می‌شود.

### ۱-۲-۳ متغیرهای تصمیم‌گیری

$x_{ijt}$ : تعداد سفرهای انجام شده از مبدأ  $i$  به مقصد  $j$  در گروه زمانی  $t$ .

### ۲-۲-۳ پارامترها

$d_{ijt}$ : تقاضای پیش‌بینی شده برای سفر از  $i$  به  $j$  در بازه زمانی  $t$ .

$c_{ijt}$ : متوسط هزینه سفر از  $i$  به  $j$  در گروه زمانی  $t$ .

$l_{ij}$ : مسافت میان  $i$  و  $j$ .

$\pi$ : احتمال دریافت انعام (۰.۹۷۳۴).

-  $\delta_{ijt}$ : نسبت سفرهایی که از مبدأ  $i$  به مقصد  $j$  در گروه زمانی  $t$ ، پرداخت انعام داشته‌اند.  
 (برای محاسبه این مقدار، نسبت تعداد سفرهایی که انعام در آن‌ها پرداخت شده‌اند، به کل سفرها محاسبه می‌گردد. برای پرداخت یا عدم پرداخت، از مدل پیش‌بینی فاز اول استفاده شده‌است)

- $r$ : مبلغ انعام متوسط (۵ دلار به ازای هر سفر).
- $C$ : هزینه ثابت مقرر شده (۴۰ دلار به ازای هر سفر).
- $K$ : محدودیت مسافت روزانه (۱۰۰۰ کیلومتر).

### ۳-۲-۳ محدودیت‌ها

- محدودیت ۱: مجموع مسافت طی شده نباید از ۱۰۰۰ کیلومتر بیشتر باشد:

$$\sum_{t=1}^T \sum_i \sum_j x_{ijt} \cdot l_{ij} \leq K \quad (1)$$

- محدودیت ۲: تعداد سفرها

$$x_{ijt} \leq d_{ijt}, \quad \forall i, j, t \quad (2)$$

- محدودیت ۳: تعداد سفرها باید غیرمنفی باشد:

$$x_{ijt} \geq 0, \quad \forall i, j, t \quad (3)$$

### ۴-۲-۳ تابع هدف

- تابع هدف: حداکثر سازی درآمد شرکت از حمل و نقل با احتساب هزینه ثابت و انعام (در صورت وجود) برای هر سفر به صورت زیر است:

$$\max \sum_{t=1}^T \sum_i \sum_j x_{ijt} \cdot (C + \pi \cdot r \cdot \delta_{ijt} + c_{ijt}) \quad (4)$$

### ۳-۳ ماتریس تقاضا

در این قسمت، پس از محاسبه ماتریس OD، یک زیرماتریس با ابعاد  $25 \times 25$  با بیشترین وزن ممکن استخراج شده است و سپس با استفاده از مدل جاذبه<sup>۱</sup> با دوتابع بازدارندگی<sup>۲</sup> متفاوت، ماتریس‌های پیش‌بینی OD سال آینده، بدست آمده‌اند.

#### ۱-۳-۳ تابع نمایی و log-normal

تابع بازدارندگی مورد استفاده به شرح زیراند:

$$f(c_{ij}) = \alpha \cdot \exp(-\beta c_{ij}) \quad (5)$$

$$f(c_{ij}) = \alpha \cdot \exp(-\beta \cdot \ln^*(c_{ij} + 1)) \quad (6)$$

تابع بازدارندگی نمایی و log-normal به ترتیب در روابط (۵) و (۶) آمده‌اند. این توابع باید به ازای مقادیر  $\alpha = ۰/۱$  و  $\beta = ۰/۱$  و بر اساس متوسط هزینه‌های بدست آمده با حداقل خطای درصد، محاسبه شوند.

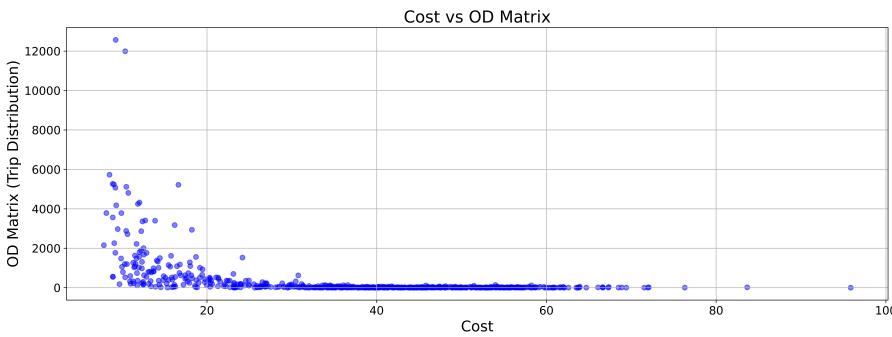
#### ۲-۳-۳ انتخاب ماتریس OD مناسب

در این تحلیل، از دو مدل مختلف برای پیش‌بینی ماتریس تقاضا استفاده شده است: مدل نمایی و مدل log-normal. نتایج نشان می‌دهند که مدل نمایی عملکرد بهتری نسبت به مدل log-normal دارد. در مدل نمایی، خطای کل برابر با  $۴۵۲/۱۵۸$  و درصد خطای برابر با  $۲/۱۱$  درصد است. همچنین، معیارهای ارزیابی مدل نمایی شامل RMSE برابر با  $۴۵۲/۱۵۸$  و  $R^2$  برابر با  $۰/۸۵$  هستند که نشان‌دهنده دقیق‌تری پیش‌بینی این مدل می‌باشد. در مقایسه، مدل log-normal دارای خطای کمتری با مقدار  $۳۴۷/۲۳$  و درصد خطای برابر با  $۰/۱۶$  درصد است، اما RMSE آن  $۵۳۵/۳۱$  و  $R^2$  آن  $۰/۷۴$  است که از دقیق‌تری نسبت به مدل نمایی حکایت دارد.

این نتایج نشان می‌دهند که مدل نمایی برای این داده‌ها مناسب‌تر است. در مدل نمایی، تعداد سفرها به طور نمایی با افزایش هزینه کاهش می‌یابند که به طور طبیعی با داده‌های دنیای

---

Gravity<sup>۱</sup>  
Deterrence Function<sup>۲</sup>



شکل ۱۱: هزینه سفر در برابر تقاضای آن

واقعی، که در آن تقاضا به طور سریع تر و با حساسیت بیشتری به تغییرات هزینه‌ها واکنش نشان می‌دهد، سازگار است. از سوی دیگر، مدل log-normal به دلیل استفاده ازتابع لگاریتمی برای مدل‌سازی هزینه‌ها، در این حالت کمتر موفق عمل کرده است.

شکل ۱۱، نشان می‌دهد که تعداد سفرها حساسیت نسبتاً بالایی به تغییرات هزینه دارند. این حساسیت بالا به ویژه در مدل نمایی مشهود است، که به خوبی تغییرات در تقاضا را با توجه به تغییرات هزینه‌ها شبیه‌سازی می‌کند. در واقع، مشاهده می‌شود که تغییرات کوچک در هزینه‌ها می‌توانند تأثیر زیادی بر تعداد سفرها بگذارند. به این ترتیب، مدل نمایی با توجه به دقت بالاتر و حساسیت به هزینه‌ها، انتخاب بهتری برای استفاده در مسئله بهینه‌سازی است.

### ۳-۳-۳ نتایج

با حل مدل با استفاده از کتابخانه pulp در زبان برنامه‌نویسی پایتون<sup>۴</sup> مقدار تابع هدف، ۷۲۸۷۲,۲۵ واحد پولی بدست می‌آید. همچنین مقادیر بهینه بسیاری از متغیرهای تصمیم‌گیری صفر بدست آمده که با توجه به محدودیت‌های شرکت مورد نظر و پارامترهای بازار، می‌توان آن را منطقی دانست. فایل متغیرها، محدودیت‌ها و تابع هدف و مقادیر بهینه آنان به ترتیب در فایل‌های model.lp و modelresults.txt آمده است.

در نهایت، تمامی فایل‌های پروژه برای فاز اول و فاز دوم، در **لينک گيت‌هاي** مربوطه موجود و قابل دسترسی است.