서울시 교통사고 지역별 통계로 알아본 사고 유형과 차량 종류

데이터프로세싱 레포트 발표

국가통계전공 2019380501 한민주

CONTENTS

01 목표

02 데이터

03 데이터 분석 및 시각화

04 결과

CONTENTS

01 목표

- 2016년부터 2020년도까지 서울시 지역별 교통사고 현황을 분석
- 사고 유형 중 차대사람 유형만, 차량 종류 중 대표적인 종류 6가지만 분석
- 이를 통해 어느 지역에서 사고가 많이 일어나는지 알아보고 그 이유는 무엇인지 예측
- 사고 이유를 예측하고 해결방안을 생각

02 데이터

데이터 출처

사고 유형별 교통사고 현황과 차종별 교통사고 현황 데이터(2016-2020년)

- 교통사고분석시스템

http://taas.koroad.or.kr/sta/acs/gus/selectTrnsportCndTfcacdSttus.do?menuId=WEB_KMP_OVT_MVT_TAC_TCT

- 시도, 시군구별 사고유형별 교통사고
- 시도, 시군구별 가해운전자 차종별 교통사고

http://taas.koroad.or.kr/sta/acs/exs/typical.do?menuId=WEB_KMP_OVT_UAS_TAT

02 데이터

데이터_사고유형별

23열 744행

시도	시군구	사고년도	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016
시도	시군구	사고년도	합계	차대사람	차대사람	차대사람	차대사람	차대사람	차대사람	차대차	차대차	차대차	차대차	차대차	차량단독	차량단독	차량단독	차량단독	차량단독	차량단독	차량단독	철길건널목	철길건널목
시도	시군구	사고년도	합계	합계	횡단중	차도통행중	길가장자리구역통행중	보도통행중	기타	합계	정면충돌	측면충돌	추돌	기타	합계	공작물충돌	주/정차차량 충돌	도로이탈	기타	전도전복	합계	기타
합계	합계	사고건수	220,917	48,489	26,823	3,622	3,544	1,627	12,873	162,009	19,498	73,397	41,023	28,091	10,413	2,160	36	263	3,100	4,854	6	6
합계	합계	사망자수	4,292	1,662	1,085	146	65	31	335	1,766	345	647	483	291	864	239	1	46	78	500	0	0
합계	합계	부상자수	331,720	49,745	27,431	3,653	3,627	1,726	13,308	269,269	34,011	116,959	76,935	41,364	12,698	2,944	41	402	3,479	5,832	8	8
서울	합계	사고건수	40,039	10,683	4,650	1,111	841	498	3,583	28,026	2,109	11,904	7,844	6,169	1,330	266	6	7	753	298	-	-
서울	합계	사망자수	348	197	115	26	6	3	47	119	27	50	23	19	32	10	0	0	6	16	-	-
서울	합계	부상자수	55,669	11,143	4,868	1,152	867	532	3,724	42,931	3,304	17,405	13,747	8,475	1,595	419	8	7	838	323	-	-
서울	종로구	사고건수	1,299	393	138	67	44	26	118	860	36	471	239	114	46	9	1	-	28	8	-	-
서울	종로구	사망자수	9	7	2	3	0	0	2	2	1	0	1	0	0	0	0	-	0	0	-	-
서울	종로구	부상자수	1,748	409	149	67	47	26	120	1,287	65	692	399	131	52	12	1	-	31	8	-	-
서울	중구	사고건수	1,288	345	145	40	25	16	119	906	16	460	263	167	37	8	-	-	22	7	-	-
서울	중구	사망자수	17	11	9	0	0	0	2	6	0	4	2	0	0	0	-	-	0	0	-	-
서울	중구	부상자수	1,716	357	147	41	25	17	127	1,321	23	673	404	221	38	8	-	-	22	8	-	-
서울	용산구	사고건수	1,428	302	148	3	31	7	113	1,079	194	358	299	228	47	13	-	-	19	15	-	-
서울	용산구	사망자수	12	8	7	0	0	0	1	3	1	0	1	1	1	0	-	-	0	1	-	-
서울	용산구	부상자수	2,069	313	153	3	31	7	119	1,704	327	525	532	320	52	16	-	-	20	16	-	-
서울	성동구	사고건수	1,099	281	120	17	20	18	106	776	49	341	182	204	42	12	-	-	15	15	-	-
서울	성동구	사망자수	15	9	4	0	1	1	3	2	1	1	0	0	4	0	-	-	2	2	-	-
서울	성동구	부상자수	1,541	283	120	17	19	18	109	1,209	80	488	356	285	49	18	-	-	14	17	-	-
서울	동대문구	사고건수	2,051	584	314	32	39	12	187	1,407	244	549	259	355	60	4	-	-	29	27	-	-
서울	동대문구	사망자수	14	8	7	0	1	0	0	5	2	0	2	1	1	0	-	-	0	1	-	-
서울	동대문구	부상자수	2,629	603	323	34	41	13	192	1,963	345	715	482	421	63	6	-	-	29	28	-	-
서울	성북구	사고건수	1,618	482	269	63	51	12	87	1,074	59	570	303	142	62	19	-	-	31	12	-	-
서울	성북구	사망자수	13	7	5	0	0	0	2	3	0	3	0	0	3	2	-	-	0	1	-	-
서울	성북구	부상자수	2,106	500	283	66	53	12	86	1,539	87	789	492	171	67	20	-	-	35	12	-	-
서울	도봉구	사고건수	949	298	141	23	18	12	104	621	48	286	174	113	30	3	-	-	20	7	-	-
서울	도봉구	사망자수	11	9	4	3	0	0	2	2	0	1	0	1	0	0	-	-	0	0	-	-
서울	도봉구	부상자수	1,306	306	145	20	19	13	109	966	81	431	295	159	34	4	-	-	22	8	-	-
서울	은평구	사고건수	1,121	394	177	40	14	27	136	685	17	356	177	135	42	11	-	1	24	6	-	-
서울	은평구	사망자수	14	8	5	0	0	0	3	4	0	2	0	2	2	2	-	0	0	0	-	-
서울	은평구	부상자수	1,518	415	184	42	14	27	148	1,047	29	554	290	174	56	16	-	1	32	7	-	-
서울	서대문구	사고건수	1,135	314	104	67	9	12	122	767	86	313	248	120	54	11	-	1	34	8	-	-
서울	서대문구	사망자수	14	11	6	2	1	0	2	3	0	2	0	1	0	0	-	0	0	0	-	-
서울	서대문구	부상자수	1,566	322	109	70	8	12	123	1,182	129	478	422	153	62	11	-	1	41	9	-	-
서울	마포구	사고건수	1,768	437	159	96	29	19	134	1,273	174	535	405	159	58	12	-	1	32	13	-	-
서울	마포구	사망자수	15	8	3	4	0	0	1	5	4	1	0	0	2	1	-	0	1	0	-	-
서울	마포구	부상자수	2,605	455	167	101	30	19	138	2,080	270	839	742	229	70	12	-	1	38	19	-	-
서울	강서구	사고건수	1,691	471	281	38	13	11	128	1,139	218	394	194	333	81	7	-	-	57	17	-	-

Windows 정품 인증
[설정]으로 이동하여 Windows를 정품 인증하십시오

02 데이터

데이터_차종별

14열 743행

시도	시군구	사고년도	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016	2016
시도	시군구	사고년도	합계	승용차	승합차	화물차	특수차	이륜차	동기장치자전	자전거	건설기계	농기계	불명
합계	합계	사고건수	220,917	147,692	14,109	26,576	951	13,076	5,906	5,936	2,482	443	3,746
합계	합계	사망자수	4,292	2,117	283	952	44	428	186	113	94	73	2
합계	합계	부상자수	331,720	228,777	23,075	40,794	1,483	15,773	6,991	6,292	3,886	506	4,143
서울	합계	사고건수	40,039	27,052	2,740	2,823	67	3,114	1,426	1,884	222	3	708
서울	합계	사망자수	348	168	35	48	3	48	18	15	13	0	0
서울	합계	부상자수	55,669	39,054	3,963	3,870	107	3,810	1,719	2,051	319	3	773
서울	종로구	사고건수	1,299	861	116	104	4	111	66	22	5	-	10
서울	종로구	사망자수	9	6	1	0	0	1	0	0	1	-	0
서울	종로구	부상자수	1,748	1,207	155	134	5	130	80	23	4	-	10
서울	중구	사고건수	1,288	904	118	92	-	95	40	6	8	-	25
서울	중구	사망자수	17	9	2	2	-	3	0	0	1	-	0
서울	중구	부상자수	1,716	1,245	158	111	-	112	47	6	7	-	30
서울	용산구	사고건수	1,428	1,055	74	64	1	117	30	48	5	-	34
서울	용산구	사망자수	12	7	2	1	0	1	1	0	0	-	0
서울	용산구	부상자수	2,069	1,600	106	78	2	149	32	55	7	-	40
서울	성동구	사고건수	1,099	707	40	103	5	99	63	63	7	-	12
서울	성동구	사망자수	15	7	0	2	1	2	1	1	1	-	0
서울	성동구	부상자수	1,541	1,039	57	149	4	126	78	69	7	-	12
서울	동대문구	사고건수	2,051	1,211	143	195	5	222	113	128	14	-	20
서울	동대문구	사망자수	14	8	2	1	0	2	0	1	0	-	0
서울	동대문구	부상자수	2,629	1,595	203	256	7	261	135	135	16	-	21
서울	성북구	사고건수	1,618	1,036	104	152	3	161	87	43	11	-	21

Windows 정품 인증

02 데이터

데이터분석과정

- 전국 데이터에서 서울시 데이터만 추출(사고유형별, 차종별 모두)
- 사고유형별 데이터에서 차대사람 데이터만 추출
- 차종별 데이터에서 대표적 6가지(승용차,승합차,화물차,이륜차,원동기장치자전거,자전거 데이터)만 추출
- 그래프로 나타내기

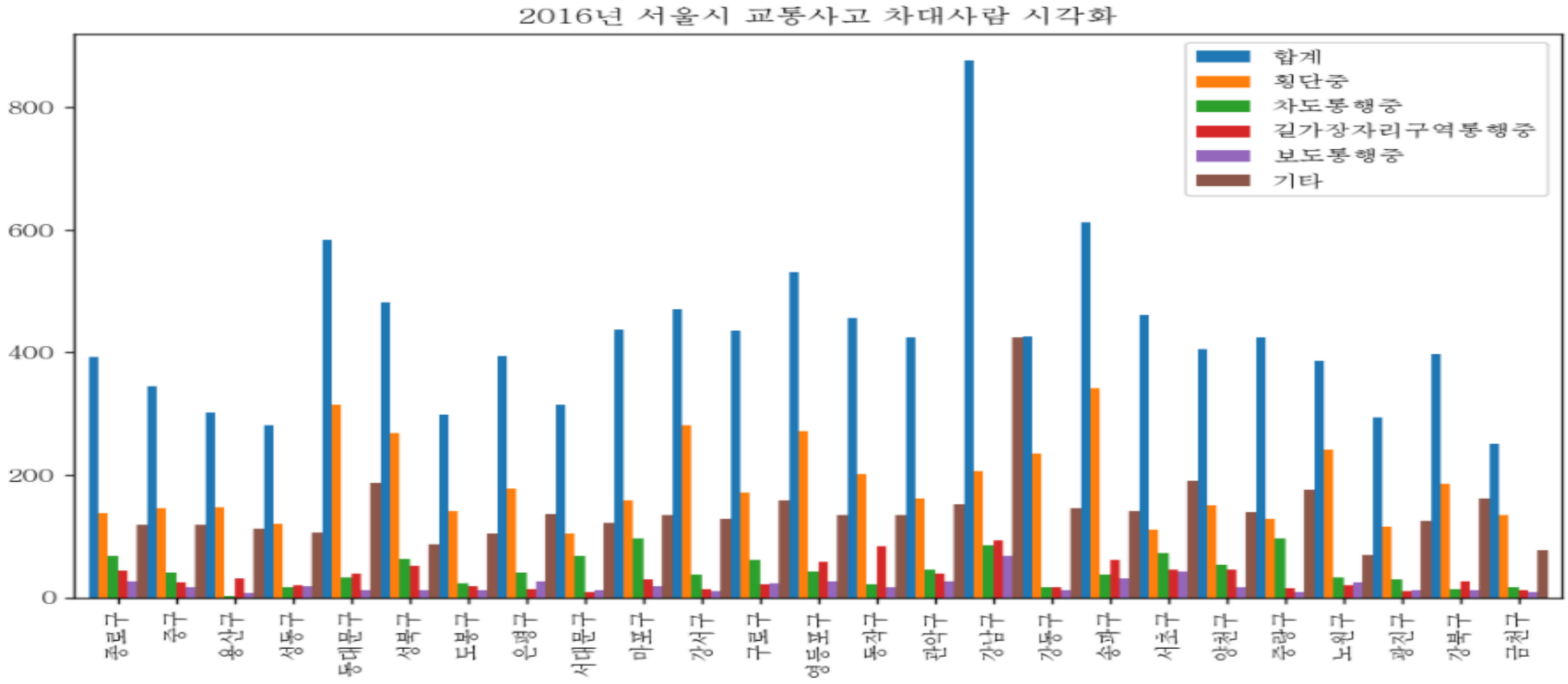
03 데이터 분석 및 시각화

사고유형별 데이터 분석 코드

```
%matplotlib notebook
from matplotlib import font_manager,rc
font_name=font_manager.FontProperties(fname="c:/Windows/Fonts/batang.ttc").get_name()
rc('font',family=font_name)
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_excel('C:/Data/type2016.xls')
df=df.query("시도=='서울' or 시도=='시도'")  #전체 데이터에서 서울시만 추출
df.drop(['시도'],axis=1,inplace=True)  #필요없는 열 제거
df=df[df["시군구"]!="합계"]   #필요한 데이터만 추출
df=df.iloc[:,0:9]   #차대사람 유형만 추출
df.set_index("시군구",inplace=True)
df.drop(['2016'],axis=1,inplace=True)
df.columns=['사고유형','합계','횡단중','차도통행중','길가장자리구역통행중','보도통행중','기타']  #칼럼 이름 변경
df=df.iloc[2:,:]   #필요없는 행 제거
df=df[df["사고유형"]=="사고건수"]   #사고건수만 추출
df.plot(kind='bar',width=1,figsize=(10,6))   #막대그래프 그리기
plt.title("2016년 서울시 교통사고 차대사람 시각화")
plt.show()
```

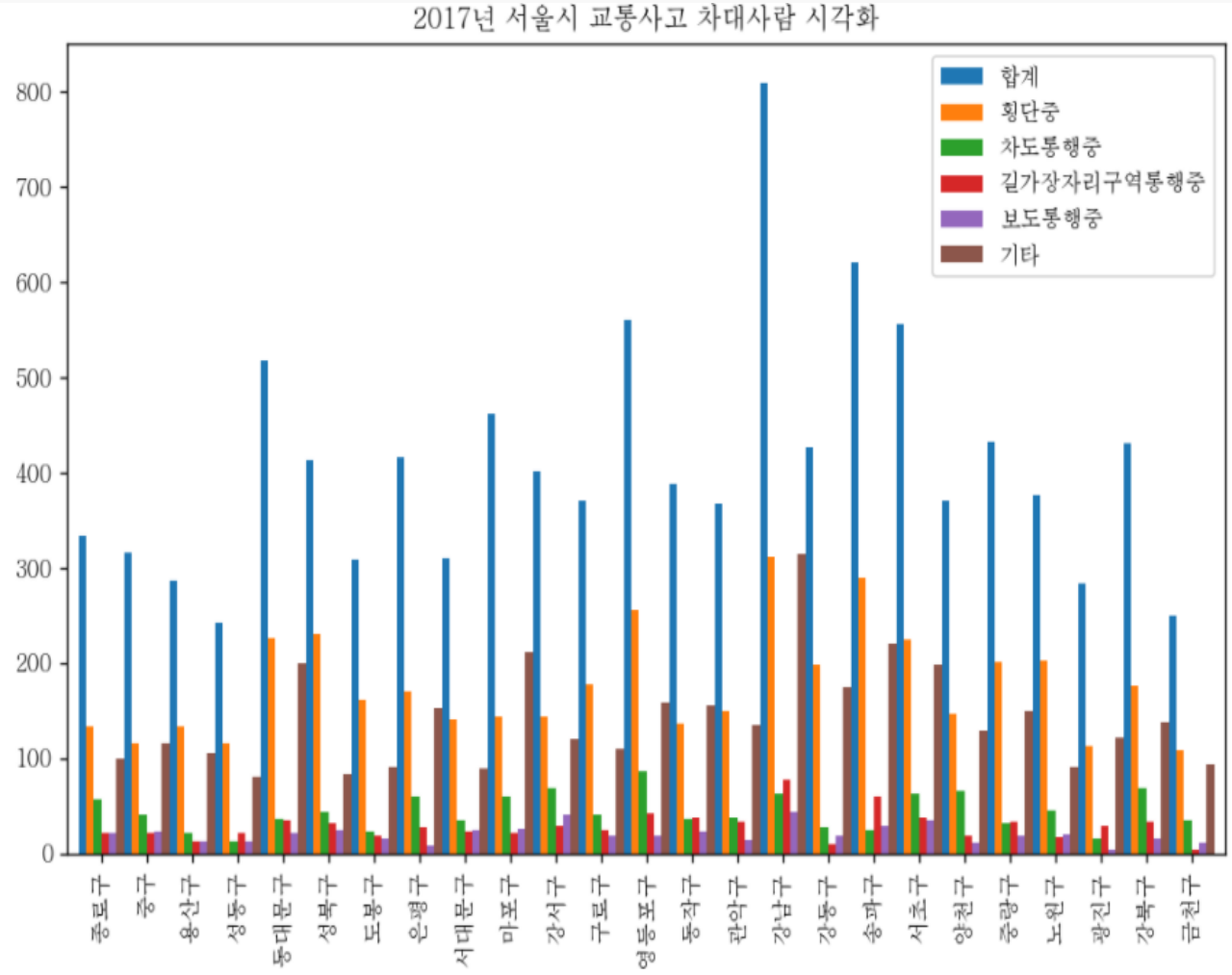
03 데이터 분석 및 시각화

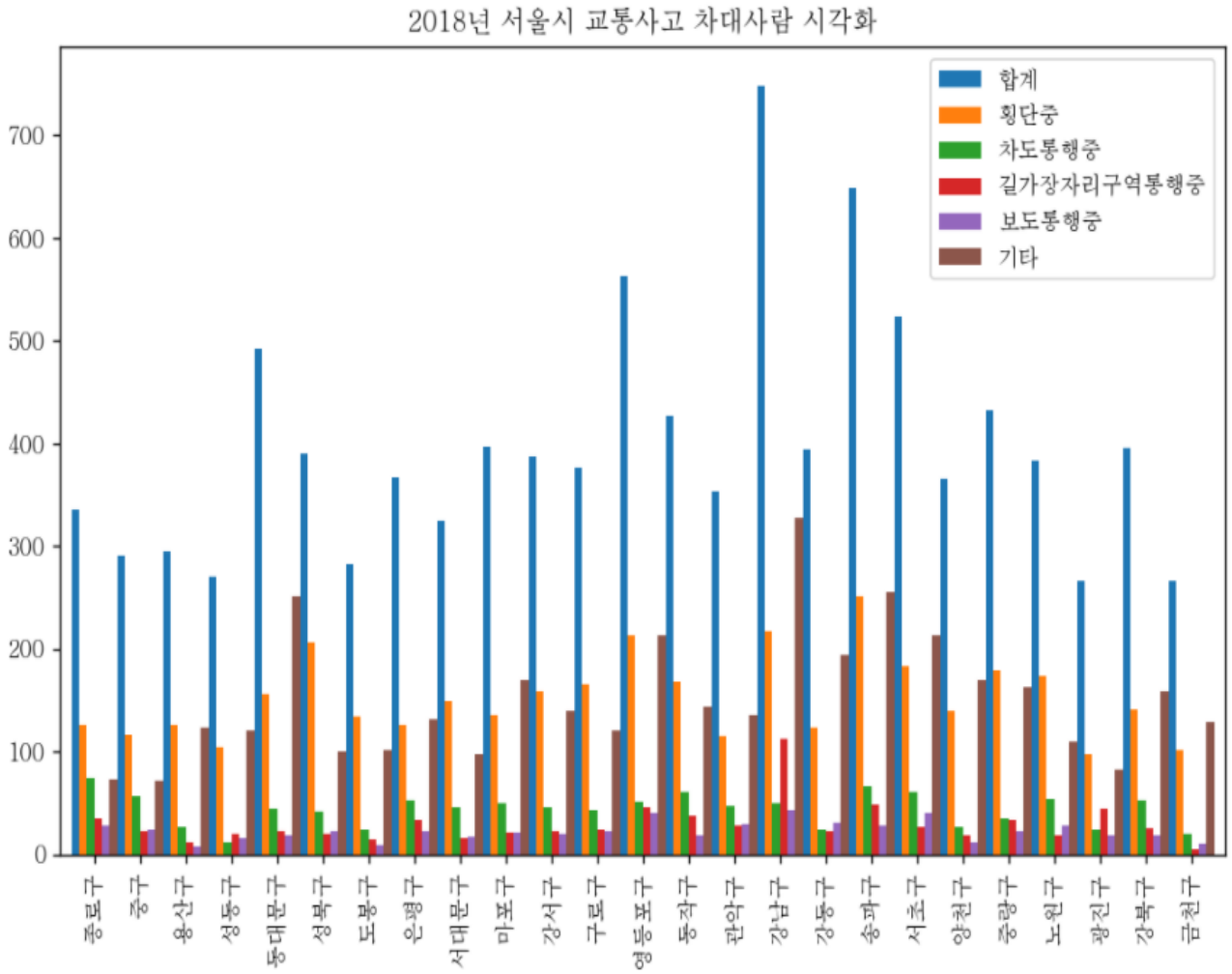
사고유형별 데이터 분석 시각화



03 데이터 분석 및 시각화

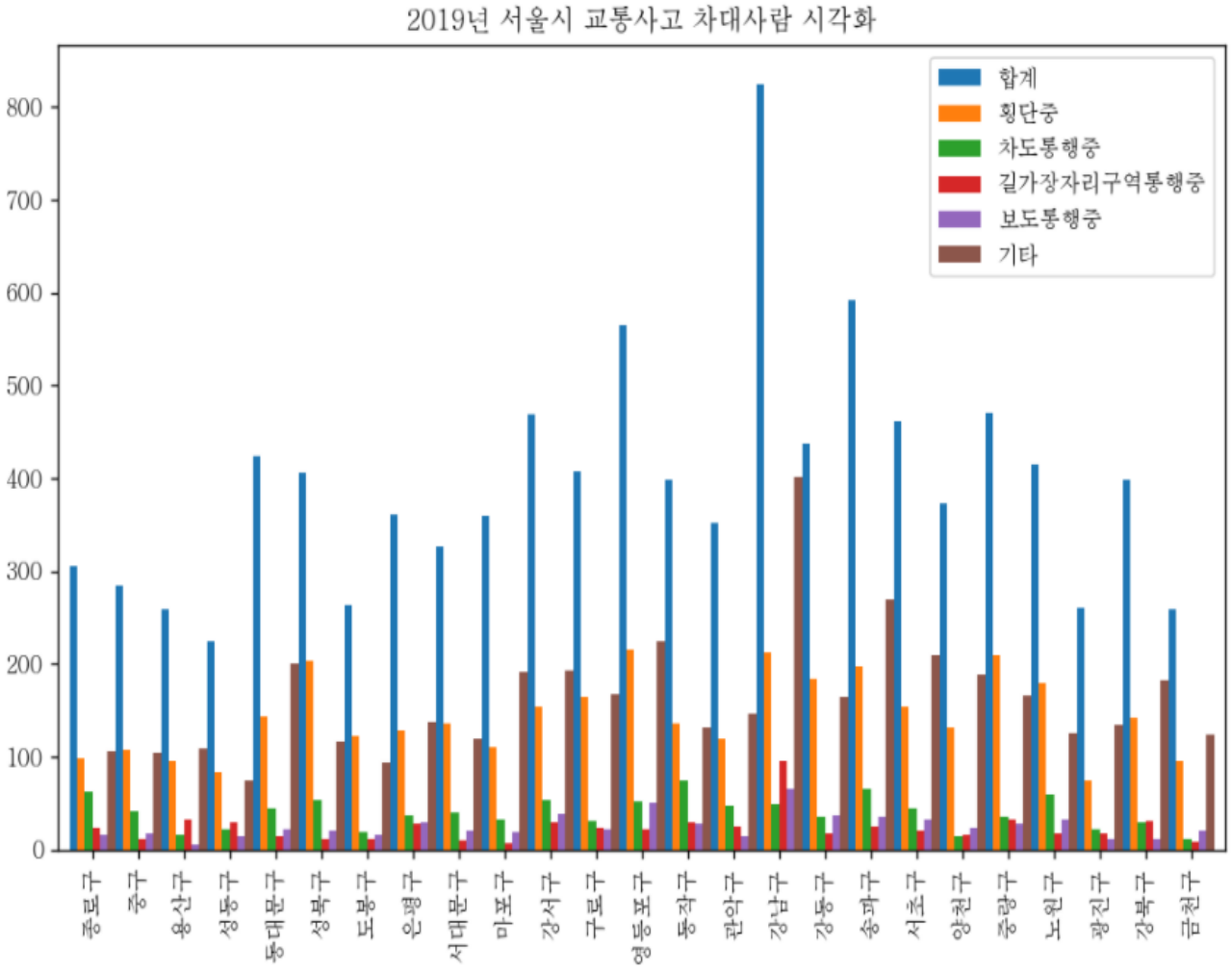
사고유형별 데이터 분석 시각화

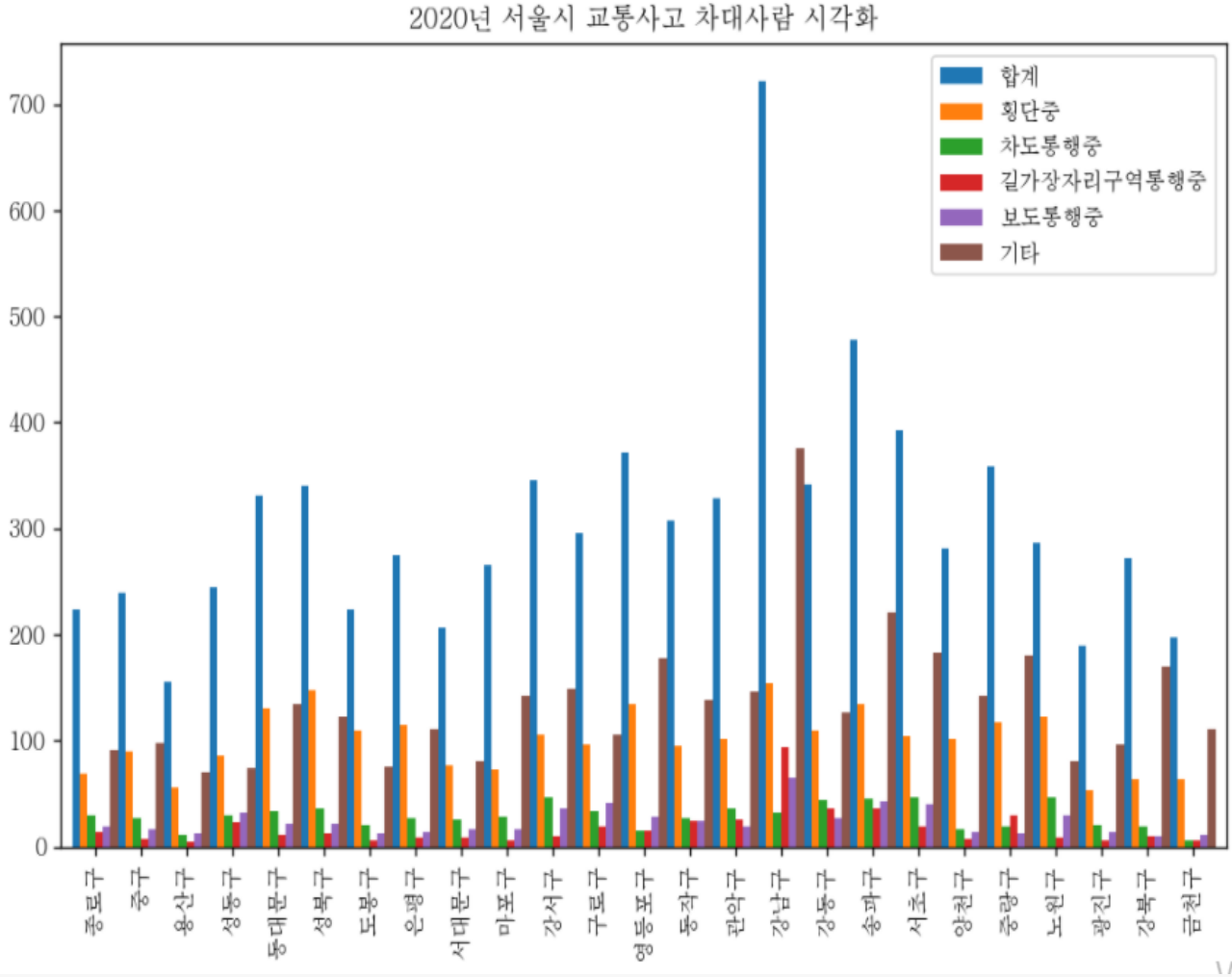




03 데이터 분석 및 시각화

사고유형별 데이터 분석 시각화





03 데이터 분석 및 시각화

사고유형별 데이터 분석-요약 통계량

```
%matplotlib notebook
from matplotlib import font_manager,rc
font_name=font_manager.FontProperties(fname="c:/Windows/Fonts/batang.ttc").get_name()
rc('font',family=font_name)
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_excel('C:/Data/type2016.xls')
df=df.query("시도=='서울' or 시도=='시도'")  #전체 데이터에서 서울시만 추출
df.drop(['시도'],axis=1,inplace=True)  #필요없는 열 제거
df=df[df["시군구"]!="합계"]   #필요한 데이터만 추출
df=df.iloc[:,0:9]   #차대사람 유형만 추출
df.set_index("시군구",inplace=True)
df.drop(['2016'],axis=1,inplace=True)
df.columns=['사고유형','합계','횡단중','차도통행중','길가장자리구역통행중','보도통행중','기타']  #칼럼 이름 변경
df=df.iloc[2:,:]   #필요없는 행 제거
df=df[df["사고유형"]=="사고건수"]   #사고건수만 추출
```

03 데이터 분석 및 시각화

사고유형별 데이터 분석-요약 통계량

```
%matplotlib notebook
from matplotlib import font_manager,rc
font_name=font_manager.FontProperties(fname="c:/Windows/Fonts/batang.ttc").get_name()
rc('font',family=font_name)
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=df.drop(df.columns[0:2],axis=1) #필요없는 열 제거
df['2016년 합계']=df.sum(axis=1)    #합계구하기
df=df.sort_values(by='2016년 합계',ascending=False)    #합계기준으로 오름차순
df1=df1.drop(df1.columns[0:2],axis=1)
df1['2017년 합계']=df1.sum(axis=1)
df1=df1.sort_values(by='2017년 합계',ascending=False)
df2=df2.drop(df2.columns[0:2],axis=1)
df2['2018년 합계']=df2.sum(axis=1)
df2=df2.sort_values(by='2018년 합계',ascending=False)
df3=df3.drop(df3.columns[0:2],axis=1)
df3['2019년 합계']=df3.sum(axis=1)
df3=df3.sort_values(by='2019년 합계',ascending=False)
df4=df4.drop(df4.columns[0:2],axis=1)
df4['2020년 합계']=df4.sum(axis=1)
df4=df4.sort_values(by='2020년 합계',ascending=False)
d=df.describe()    #2016년 요약통계량
d1=df1.describe()
d2=df2.describe()
d3=df3.describe()
d4=df4.describe()
p=pd.concat([d,d1,d2,d3,d4],axis=1)    #2016-2020년도 요약통계량 병합
p
```

03 데이터 분석 및 시각화

사고유형별 데이터 분석–요약 통계량

합계

평균

표준편차

최소값

1사분위수

2사분위수

3사분위수

최대값

	2016년 합계	2017년 합계	2018년 합계	2019년 합계	2020년 합계
count	25.000000	25.000000	25.000000	25.000000	25.000000
mean	427.320000	409.960000	399.120000	395.840000	306.840000
std	129.950542	127.626891	119.129593	129.551367	112.972593
min	251.000000	242.000000	266.000000	225.000000	155.000000
25%	345.000000	316.000000	325.000000	305.000000	239.000000
50%	424.000000	388.000000	383.000000	398.000000	286.000000
75%	461.000000	432.000000	427.000000	437.000000	342.000000
max	876.000000	810.000000	749.000000	825.000000	721.000000

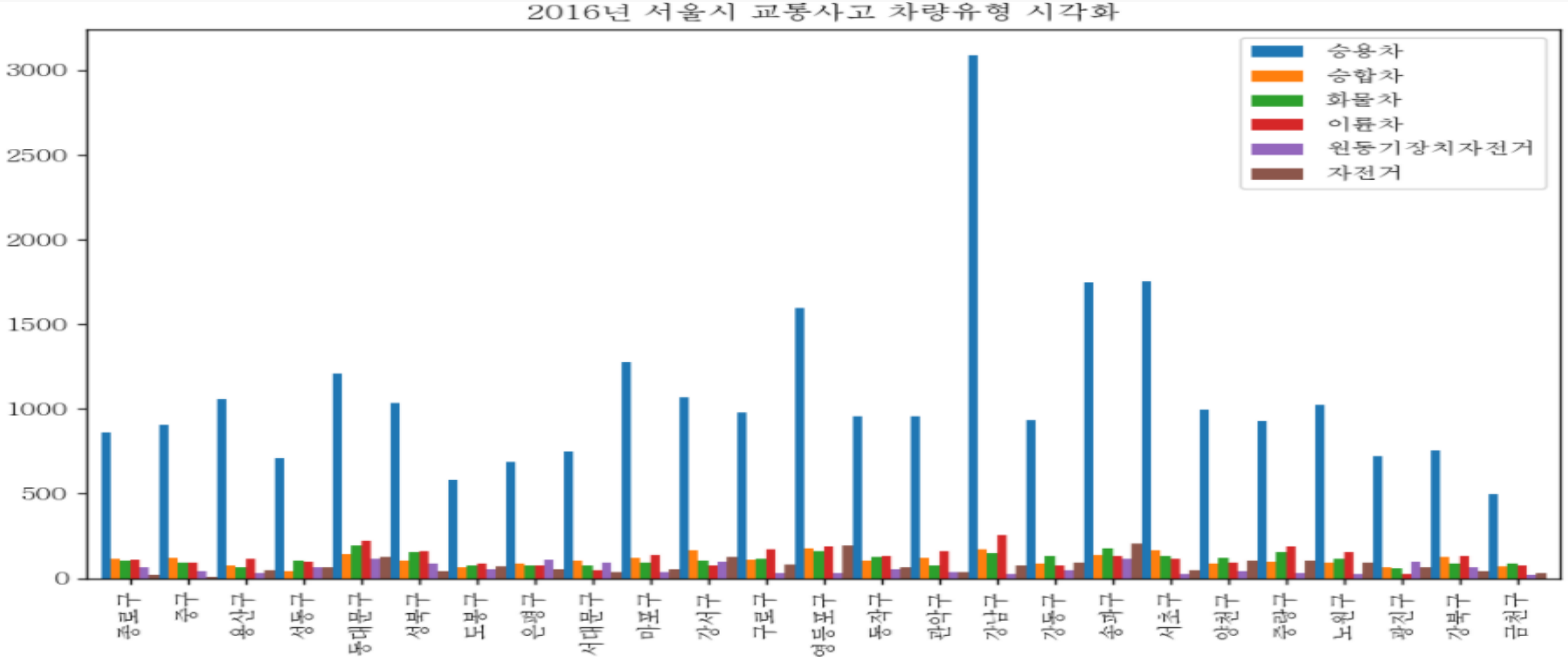
03 데이터 분석 및 시각화

차종별 데이터 분석 코드

```
%matplotlib notebook
from matplotlib import font_manager,rc
font_name=font_manager.FontProperties(fname="c:/Windows/Fonts/batang.ttc").get_name()
rc('font',family=font_name)
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_excel('C:/Data/car2016.xls')
df=df.query("시도=='서울' or 시도=='시도'")  #전체 데이터에서 서울시만 추출
df.drop(['시도'],axis=1,inplace=True)  #필요없는 열 제거
df=df[df["시군구"]!="합계"]   #필요한 데이터만 추출
df.set_index("시군구",inplace=True)
df.drop(['2016'],axis=1,inplace=True)
df=df.iloc[:,0:8]
df.drop(['2016.4'],axis=1,inplace=True)
df.columns=['차량종류','승용차','승합차','화물차','이륜차','원동기장치자전거','자전거']
df=df.iloc[1:,:]
df=df[df["차량종류"]=="사고건수"]   #사고건수만 추출
df.plot(kind='bar',width=1,figsize=(10,6))   #막대그래프 그리기
plt.title("2016년 서울시 교통사고 차량유형 시각화")
plt.show()
```

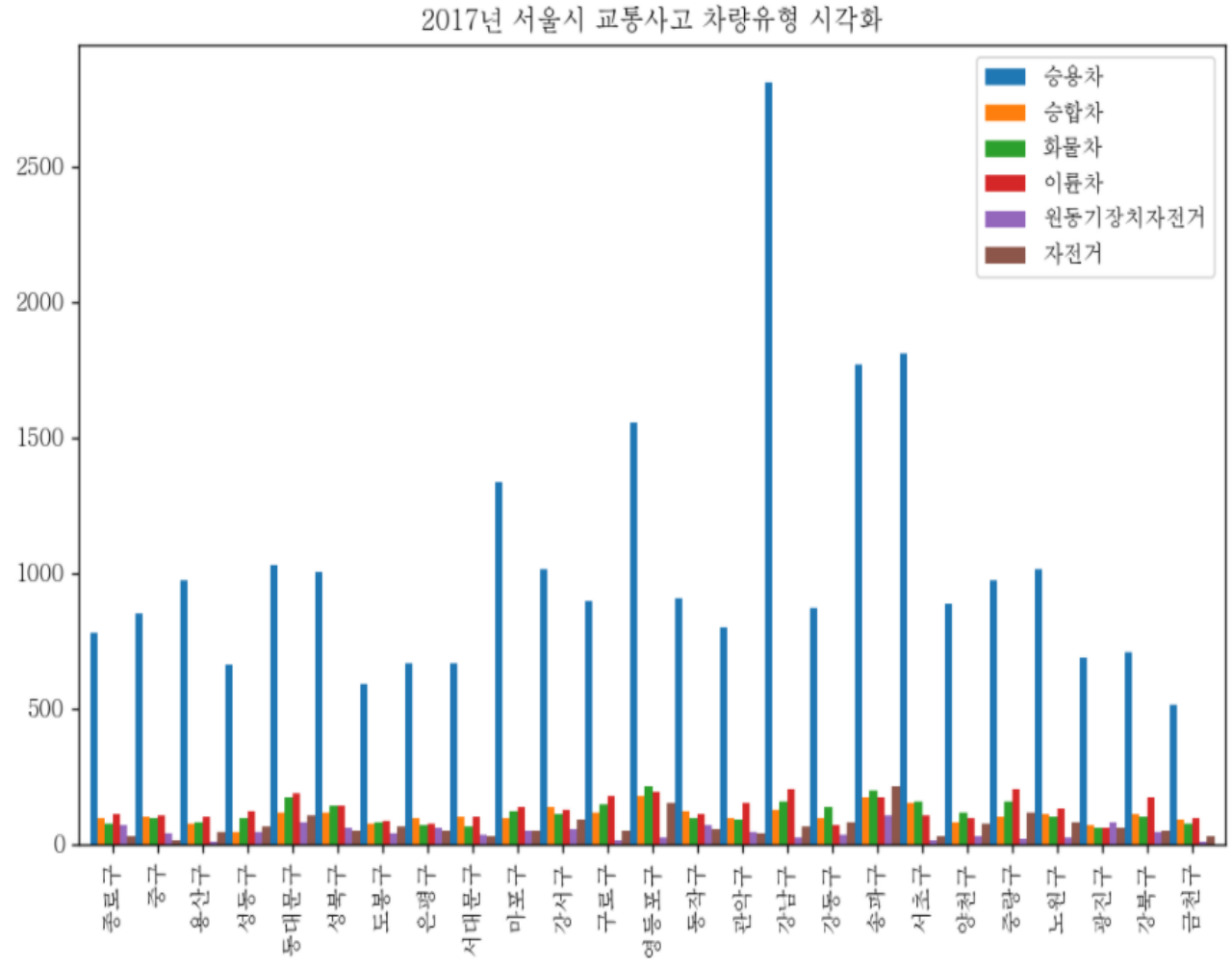
03 데이터 분석 및 시각화

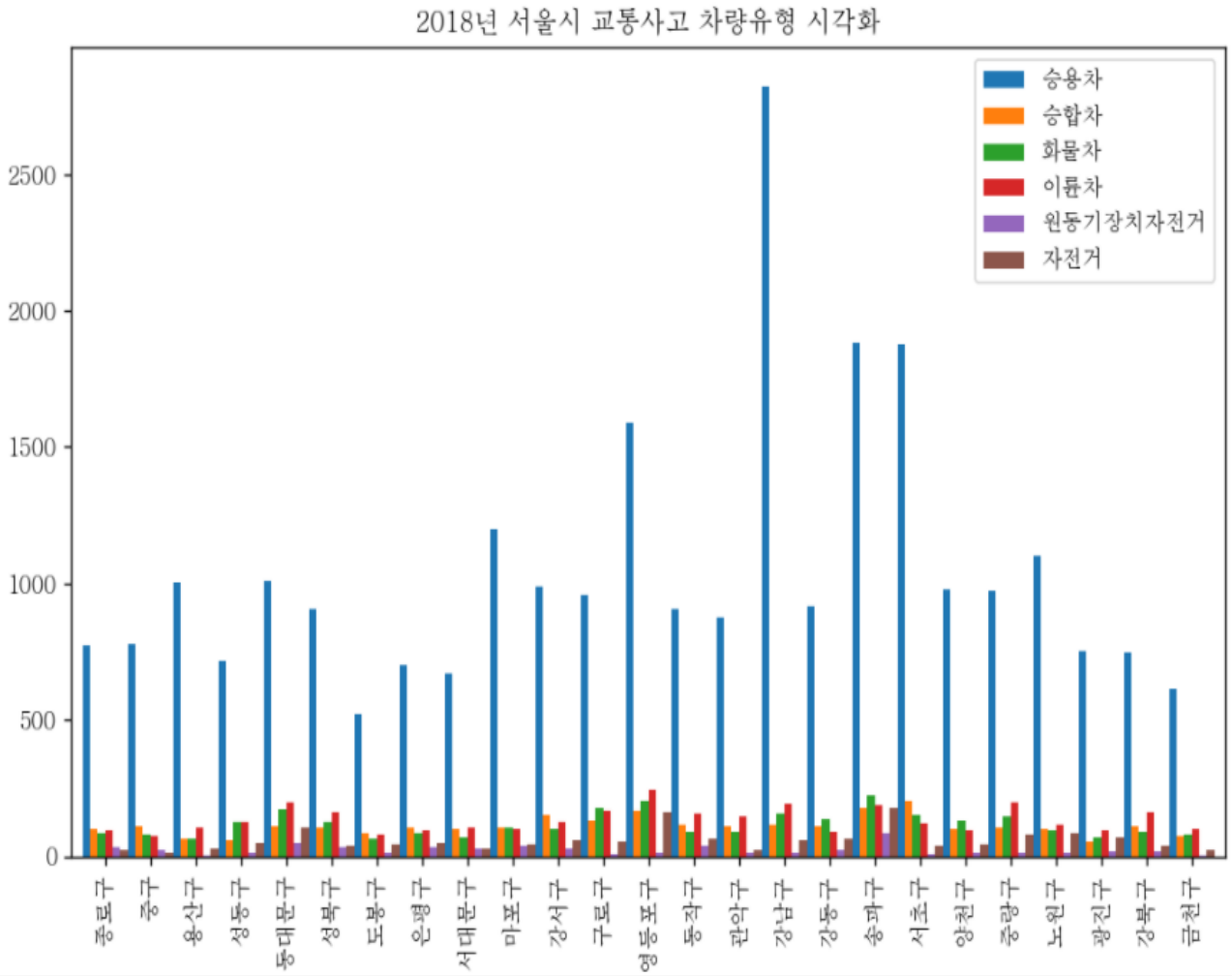
차종별 데이터 분석 시각화



03 데이터 분석 및 시각화

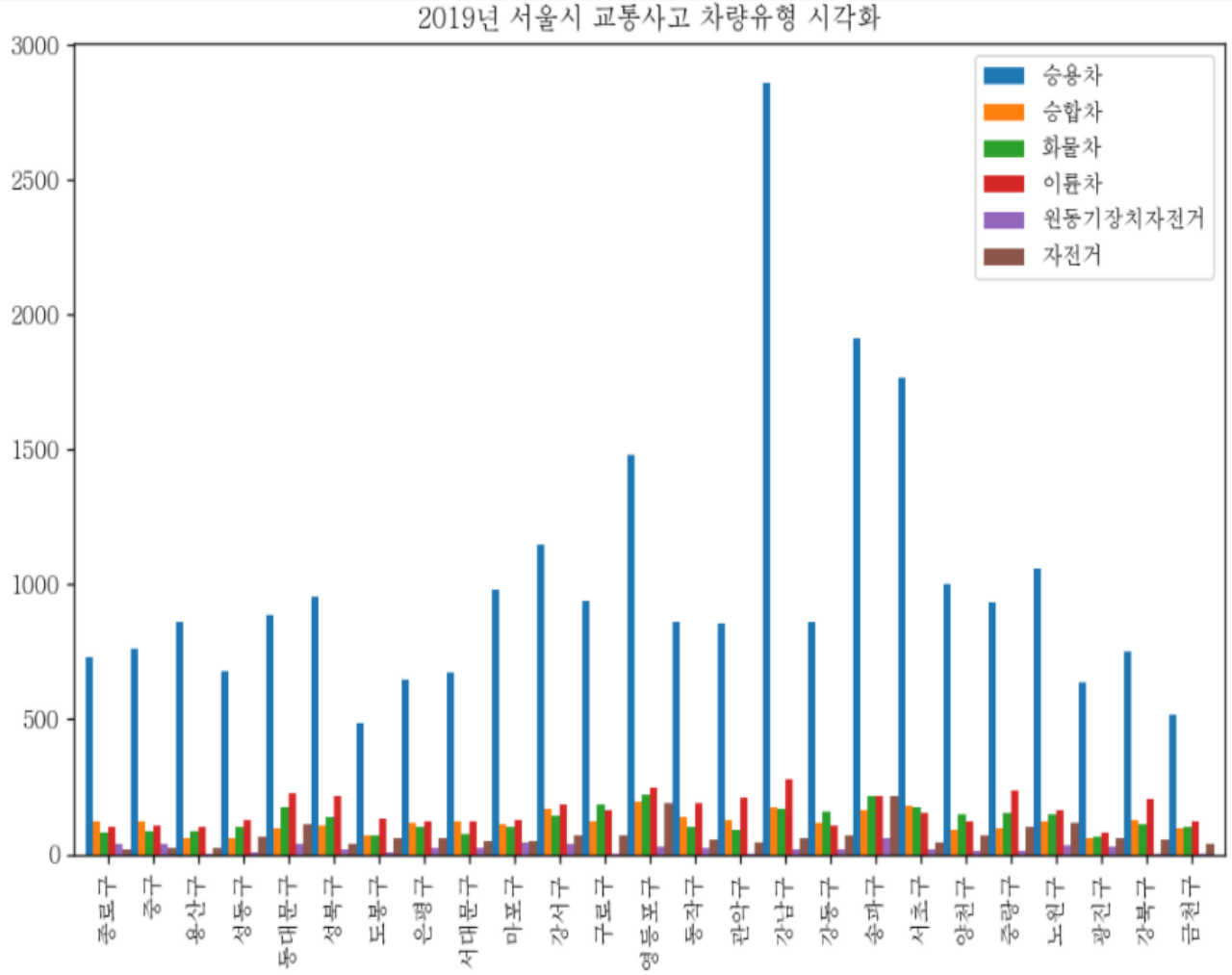
차종별 데이터 분석 시각화

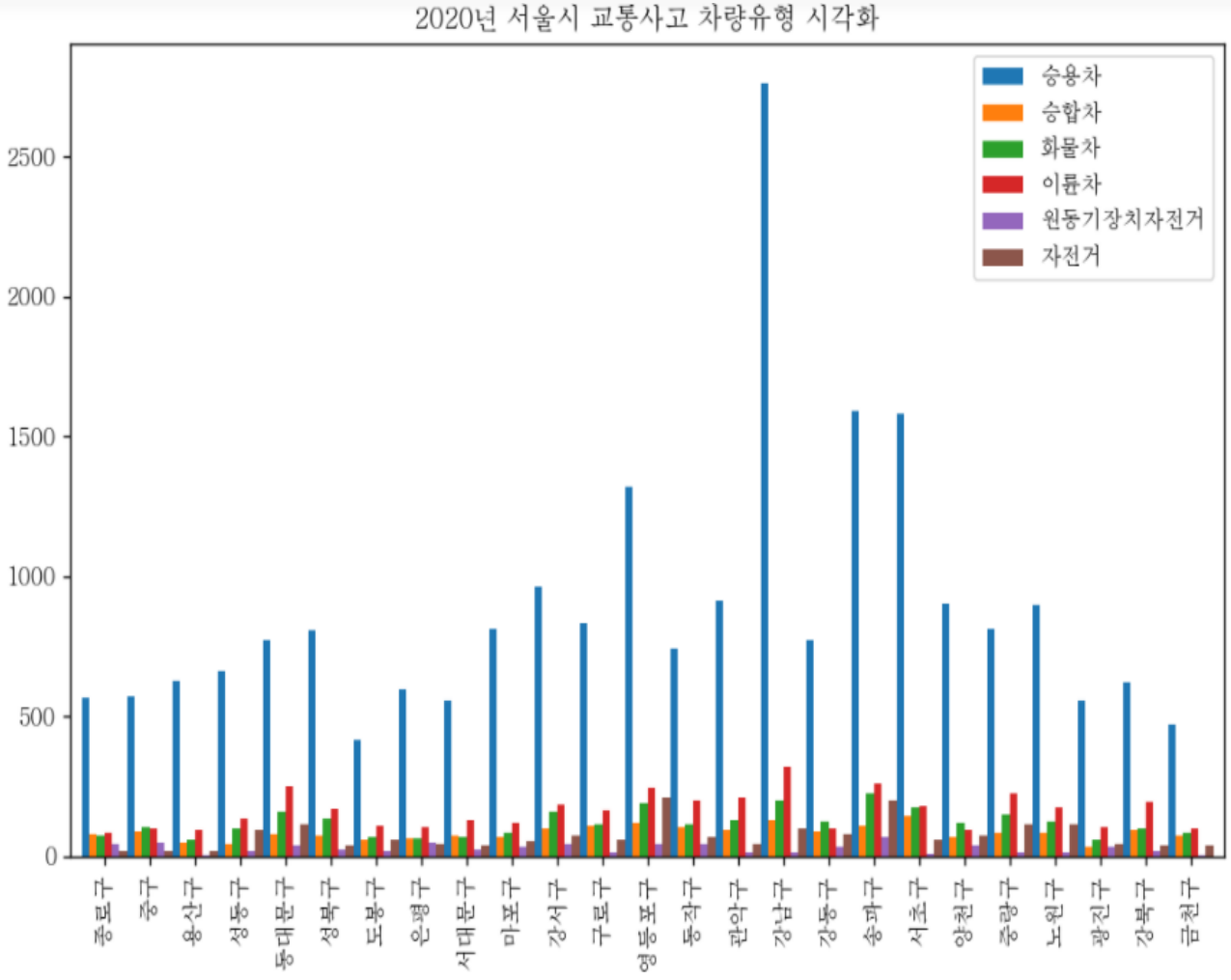




03 데이터 분석 및 시각화

차종별 데이터 분석 시각화





03 데이터 분석 및 시각화

차종별 데이터 분석-요약 통계량

```
%matplotlib notebook
from matplotlib import font_manager,rc
font_name=font_manager.FontProperties(fname="c:/Windows/Fonts/batang.ttc").get_name()
rc('font',family=font_name)
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_excel('C:/Data/car2016.xls')
df=df.query("시도=='서울' or 시도=='시도'")  #전체 데이터에서 서울시만 추출
df.drop(['시도'],axis=1,inplace=True)  #필요없는 열 제거
df=df[df["시군구"]!="합계"]   #필요한 데이터만 추출
df.set_index("시군구",inplace=True)
df.drop(['2016'],axis=1,inplace=True)
df=df.iloc[:,0:8]
df.drop(['2016.4'],axis=1,inplace=True)
df.columns=['차량종류','승용차','승합차','화물차','이륜차','원동기장치자전거','자전거']
df=df.iloc[1:,:]
df=df[df["차량종류"]=="사고건수"]   #사고건수만 추출
```

03 데이터 분석 및 시각화

차종별 데이터 분석-요약 통계량

```
%matplotlib notebook
from matplotlib import font_manager,rc
font_name=font_manager.FontProperties(fname="c:/Windows/Fonts/batang.ttc").get_name()
rc('font',family=font_name)
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df.drop(columns=['차량종류'],inplace=True)    #필요없는 열 제거
df['2016년 합계']=df.sum(axis=1)    #합계구하기
df=df.sort_values(by='2016년 합계',ascending=False)    #합계기준으로 오름차순
df1.drop(columns=['차량종류'],inplace=True)
df1['2017년 합계']=df1.sum(axis=1)
df1=df1.sort_values(by='2017년 합계',ascending=False)
df2.drop(columns=['차량종류'],inplace=True)
df2['2018년 합계']=df2.sum(axis=1)
df2=df2.sort_values(by='2018년 합계',ascending=False)
df3.drop(columns=['차량종류'],inplace=True)
df3['2019년 합계']=df3.sum(axis=1)
df3=df3.sort_values(by='2019년 합계',ascending=False)
df4.drop(columns=['차량종류'],inplace=True)
df4['2020년 합계']=df4.sum(axis=1)
df4=df4.sort_values(by='2020년 합계',ascending=False)
d=df.describe()    #2016년 요약통계량
d1=df1.describe()
d2=df2.describe()
d3=df3.describe()
d4=df4.describe()
p=pd.concat([d,d1,d2,d3,d4],axis=1)    #2016-2020년도 요약통계량 병합
p
```

03 데이터 분석 및 시각화

차종별 데이터 분석–요약 통계량

합계

평균

표준편차

최소값

1사분위수

2사분위수

3사분위수

최대값

	2016년 합계	2017년 합계	2018년 합계	2019년 합계	2020년 합계
count	25.000000	25.000000	25.000000	25.00000	25.000000
mean	1561.560000	1485.840000	1494.120000	1512.24000	1345.680000
std	627.695924	594.786144	608.471741	638.02797	630.667287
min	782.000000	801.000000	802.000000	827.00000	719.000000
25%	1206.000000	1151.000000	1086.000000	1087.00000	914.000000
50%	1433.000000	1286.000000	1365.000000	1369.00000	1239.000000
75%	1633.000000	1566.000000	1517.000000	1533.00000	1401.000000
max	3759.000000	3375.000000	3363.000000	3565.00000	3523.000000

04 결과

- 차대사람 유형은 강남구에서 가장 많이 발생(2016-2020)

=>강남구에 거주하는 사람이 많기 때문

=>2021년에도 강남구에서 가장 많이 발생할 것으로 예상

- 2016년 차대사람 교통사고 사고건수가 5년 중 가장 많음

=>갈수록 교통사고 방지를 위한 대책이 마련되고 있기 때문이라 예측

- 강남구에서 차대사람 중 기타와 횡단중 유형이 가장 많은 비중을 차지함

=>강남구에 사고가 많이 발생하는 곳은 표시하고 횡단을 할 때 조심해야함

04 결과

- 차량유형도 강남구에서 가장 많이 발생

=>강남구에 거주하는 사람이 많기 때문

- 2016년에 차량유형별 교통사고 사고건수가 5년 중 가장 많음

=> 갈수록 교통사고 방지를 위한 대책이 마련되고 있기 때문이라 예측

- 강남구에서 차량유형 중 승용차가 압도적으로 가장 많은 비중을 차지함

=>많은 사람들이 승용차를 타고 다니기 때문

=>승용차를 운전할 때 주의해야함

THANK YOU

감사합니다