

자료 파일명: 상가업소정보\_201912\_01.csv => df 변수에 저장

### ※ 결측치가 있는 컬럼 시각화하기

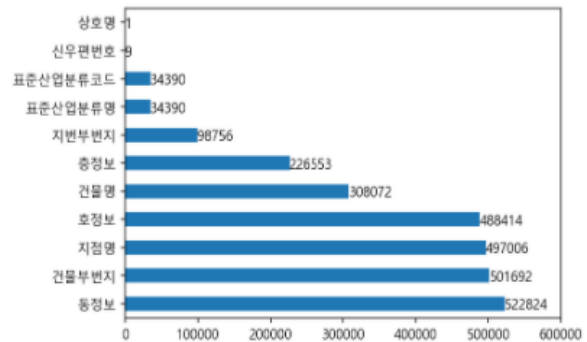
```
missing_cnt = df.isnull().sum() # df의 컬럼들 전체의 결측치 수 확인 => missing_cnt 변수에 담기
missing_cnt = missing_cnt[missing_cnt != 0] # 다시 결측치가 있는 컬럼만 missing_cnt 변수에 담기
```

## 1. barh 플롯으로 시각화

```
missing_cnt.sort_values(ascending=False)
# 내림차순 정렬

missing_cnt.plot.barh(xlim=[0, 600000])
# xlim = : x 축값 범위 지정

for i, v in enumerate(missing_cnt):
    plt.text(v, i-0.2, v) # 각 bar에 해당 수치 넣기
```



## 2. missingno 모듈을 이용하여 시각화

```
import missingno as msno
```

```
msno.matrix(df) # df 전체 컬럼의 결측치를 matrix 로 시각화 (결측치가 없는부분은 검정색)
```

