In [10]:
```
# 필요한 라이브러리 import
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import axes3d
import seaborn as sns

from sklearn.preprocessing import scale
import sklearn.linear_model as skl_lm
from sklearn.metrics import mean_squared_error, r2_score
import statsmodels.api as sm
import statsmodels.formula.api as smf
```

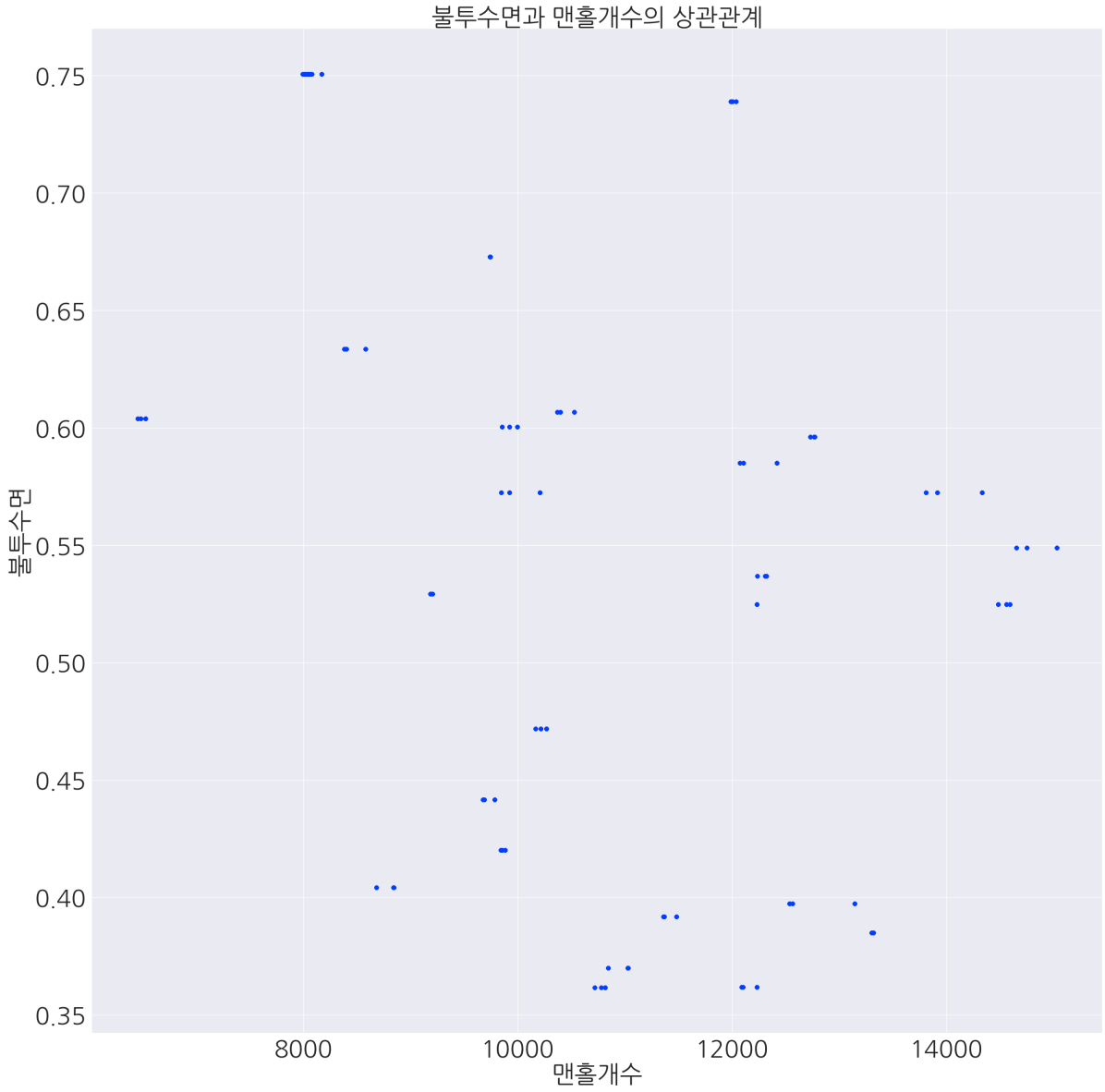
In [16]:
```
df = pd.read_excel("main.xlsx")
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11500 entries, 0 to 11499
Data columns (total 14 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   자치구(구)      11500 non-null  object 
 1   날짜          11500 non-null  int64  
 2   1hr 최대 강수량  11500 non-null  float64
 3   일평균 강수량     11500 non-null  float64
 4   경사도         11500 non-null  float64
 5   고도(해발고도)    11500 non-null  float64
 6   불투수면        11500 non-null  float64
 7   녹지 면적율      11500 non-null  float64
 8   하천 면적율      11500 non-null  float64
 9   복개하천 개수     11500 non-null  int64  
 10  맨홀개수        11500 non-null  int64  
 11  빗물받이 개수     11500 non-null  int64  
 12  빗물 펌프 개수    11500 non-null  int64  
 13  하수관로 비율     11500 non-null  object 
dtypes: float64(7), int64(5), object(2)
memory usage: 1.2+ MB
```

In [20]:
```
## 시각화
fig = plt.figure(figsize=(30,30))
fig.set_facecolor('white')
plt.rcParams['font.family'] = 'NanumGothic'
plt.title('불투수면과 맨홀개수의 상관관계', fontsize=40)

font_size = 40
plt.scatter(x=df['맨홀개수'], y=df['불투수면'])

plt.xlabel('맨홀개수', fontsize=font_size)
plt.ylabel('불투수면',fontsize=font_size)
plt.xticks(fontsize=40)
plt.yticks(fontsize=40)
plt.show()
```


불투수면과 맨홀개수의 상관관계
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
불투수면
8000
10000
12000
14000
맨홀개수

In [ ]: