

---

# Data Science

## 직종들의 연봉 분석

2조 박지수, 박상인<sub>(발표)</sub>

---



# 목차

- 데이터 정보
- 데이터 전처리
- Column 설명
- 경력별 연봉분포
- 회사 규모에 따른 연봉 차이
- 원격근로 여부별 연봉 분포
- 직무별 연봉 분포



어떻게 하면 연봉을  
더 받을 수 있을까??

# - 데이터 정보

The screenshot shows the Kaggle website interface. On the left is a navigation sidebar with the Kaggle logo and links for 'Create', 'Home', 'Competitions', 'Datasets' (which is highlighted), and 'Models'. The main content area features a search bar at the top. Below it, the dataset is attributed to 'HENRY SHAN · UPDATED 5 DAYS AGO'. The dataset title is '2023 Data Scientists Salary' with a subtitle 'Salaries of Different Data Science Fields in the Data Science Domain'. To the right of the title are buttons for 'New Notebook' and 'Download (26 kB)', along with a version indicator '5'. A small image of a mug with the text 'I NEVER DREAMED I WOULD BE A DATA SCIENTIST BUT HERE I AM KILLIN' IT' is displayed on the right side of the dataset preview.

데이터 출처

*kaggle.com*

## – 데이터 전처리

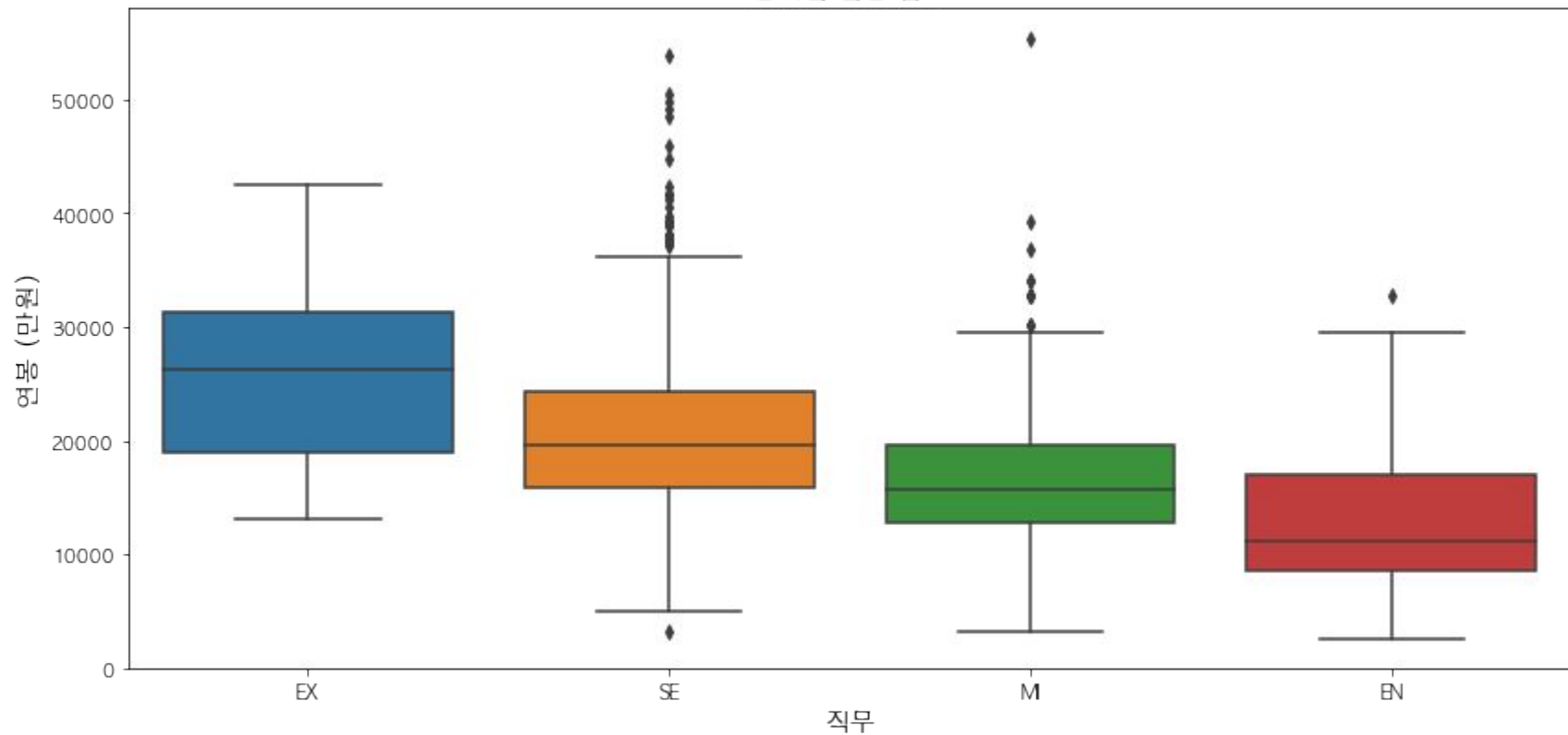
```
# 데이터 직무별 연봉 데이터 (출처: Kaggle)
df = pd.read_csv('data/ds_salaries.csv')
df = df[df['company_location'] == 'US'] # 미국 데이터만 추출
df = df[df['salary_in_usd'] > 15000] # 연봉 15000달러 이하인 데이터는 이상치로 판단하여 제거
df = df[df['employment_type'] == 'FT'] # Full-Time 근무자만 추출
df['salary_in_ko'] = df['salary_in_usd'] * 1309.86 / 10000 # 연봉 달러를 원화로 변환
```

## — Column 설명

- 
- `experience_level` - 경력 (EX, SE, MI, EN)
  - `job_title` - 직무
  - `salary_in_usd` - 연봉(달러)
  - `employee_residence` - 국가(직원)
  - `remote_ratio` - 원격 근무 비율
  - `company_location` - 국가(회사)
  - `company_size` - 회사 규모 (L/S/M)

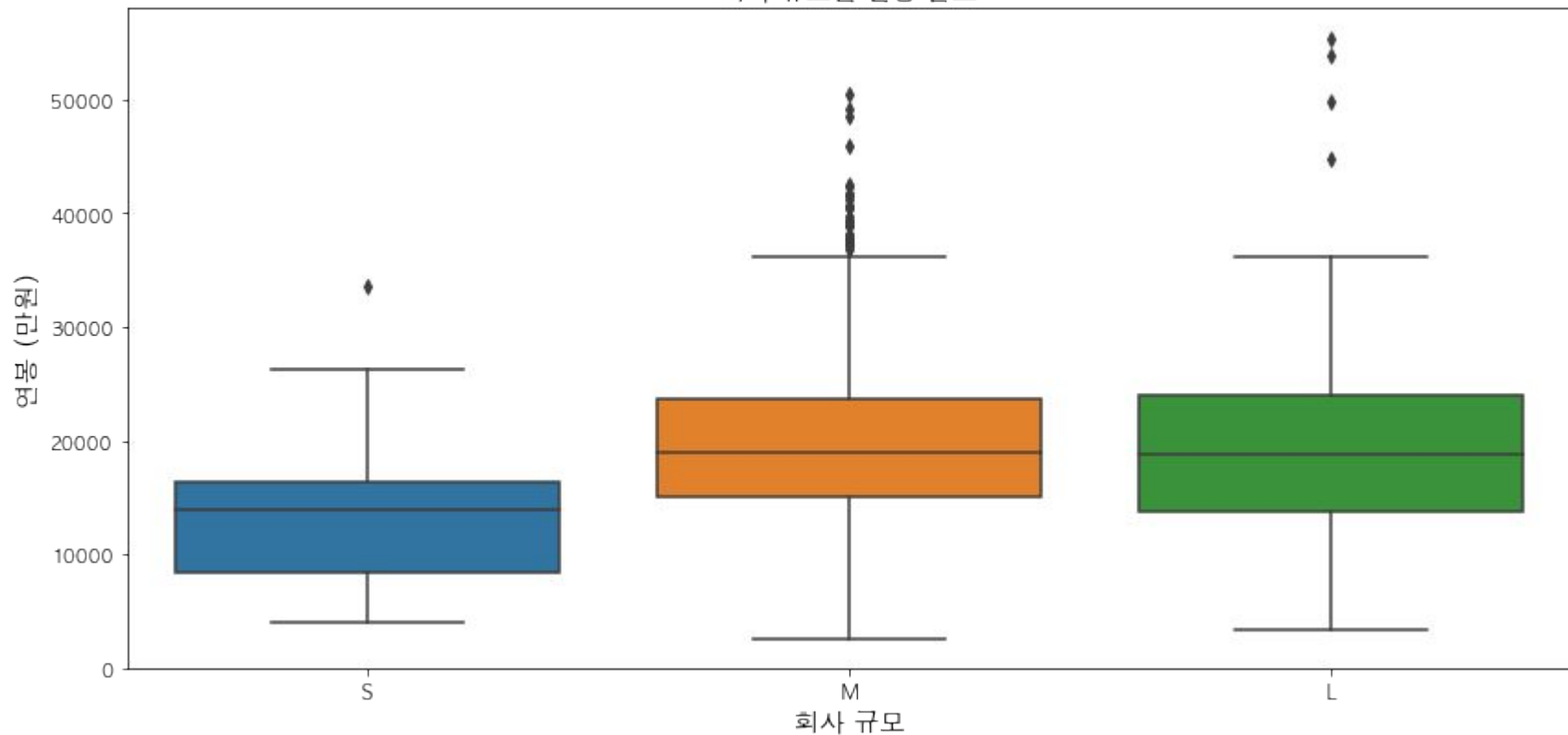
# - 경력별 연봉 분포

경력별 연봉 분포



# - 회사 규모에 따른 연봉 분포

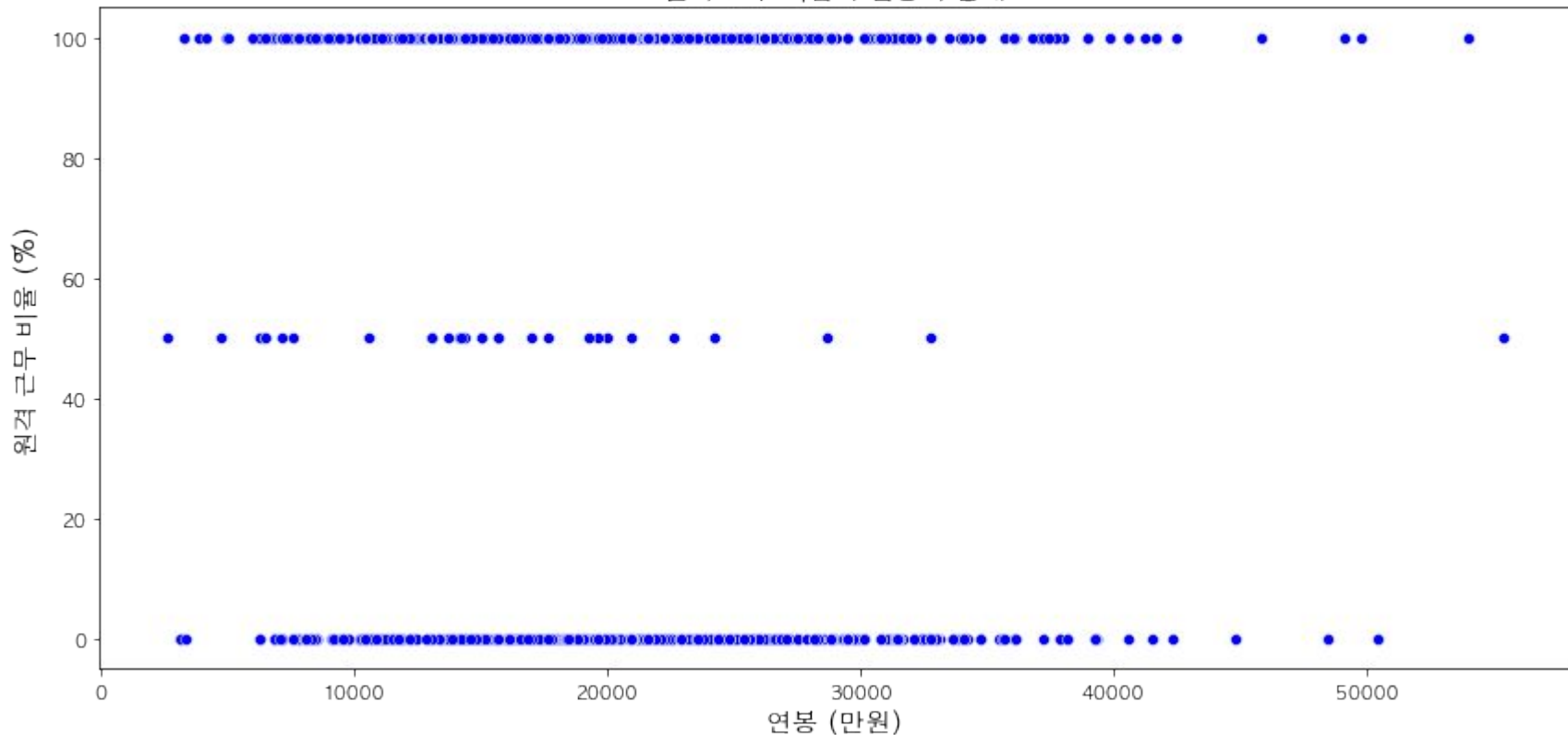
회사 규모별 연봉 분포





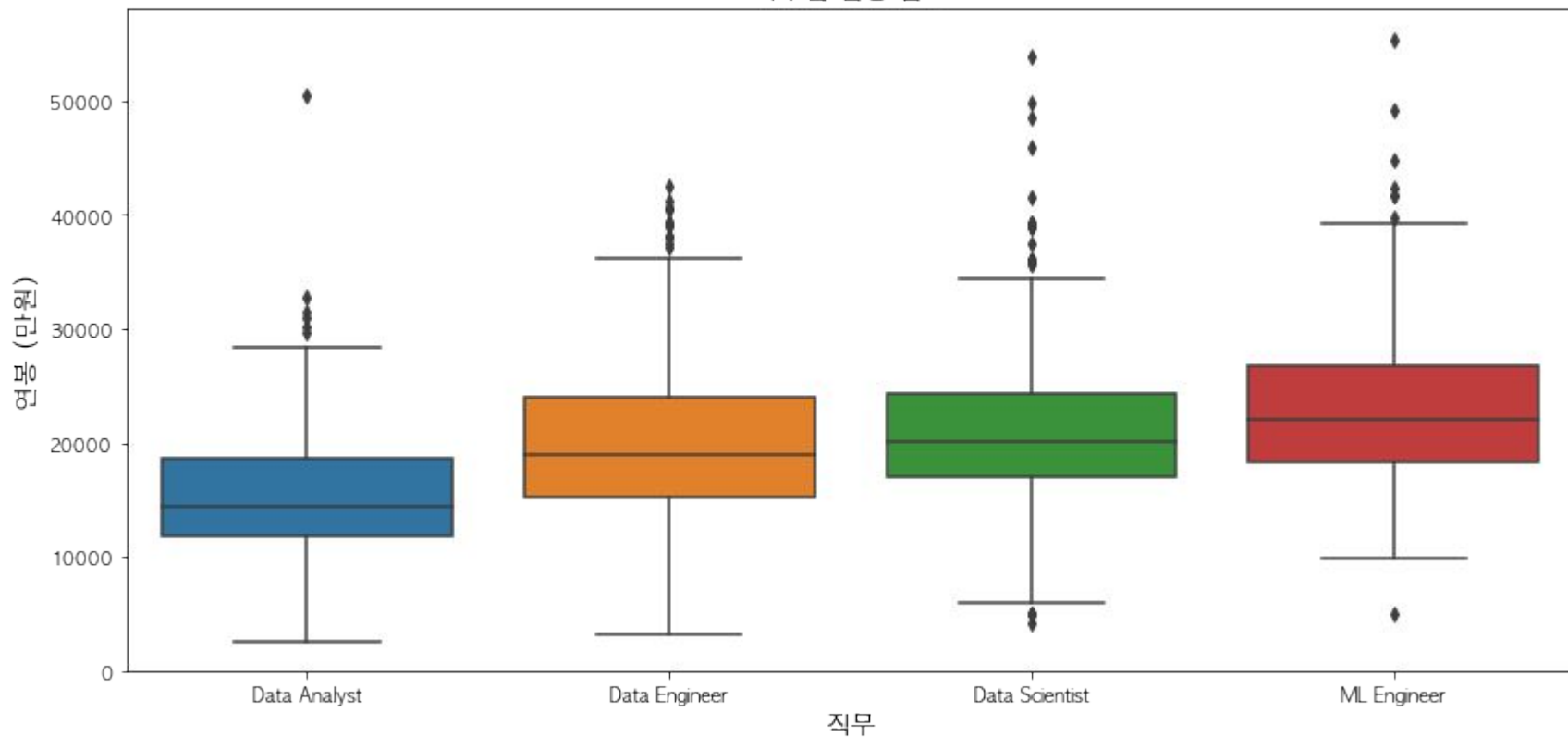
# 원격 근무 여부별 직무 분포

원격 근무 비율과 연봉의 관계

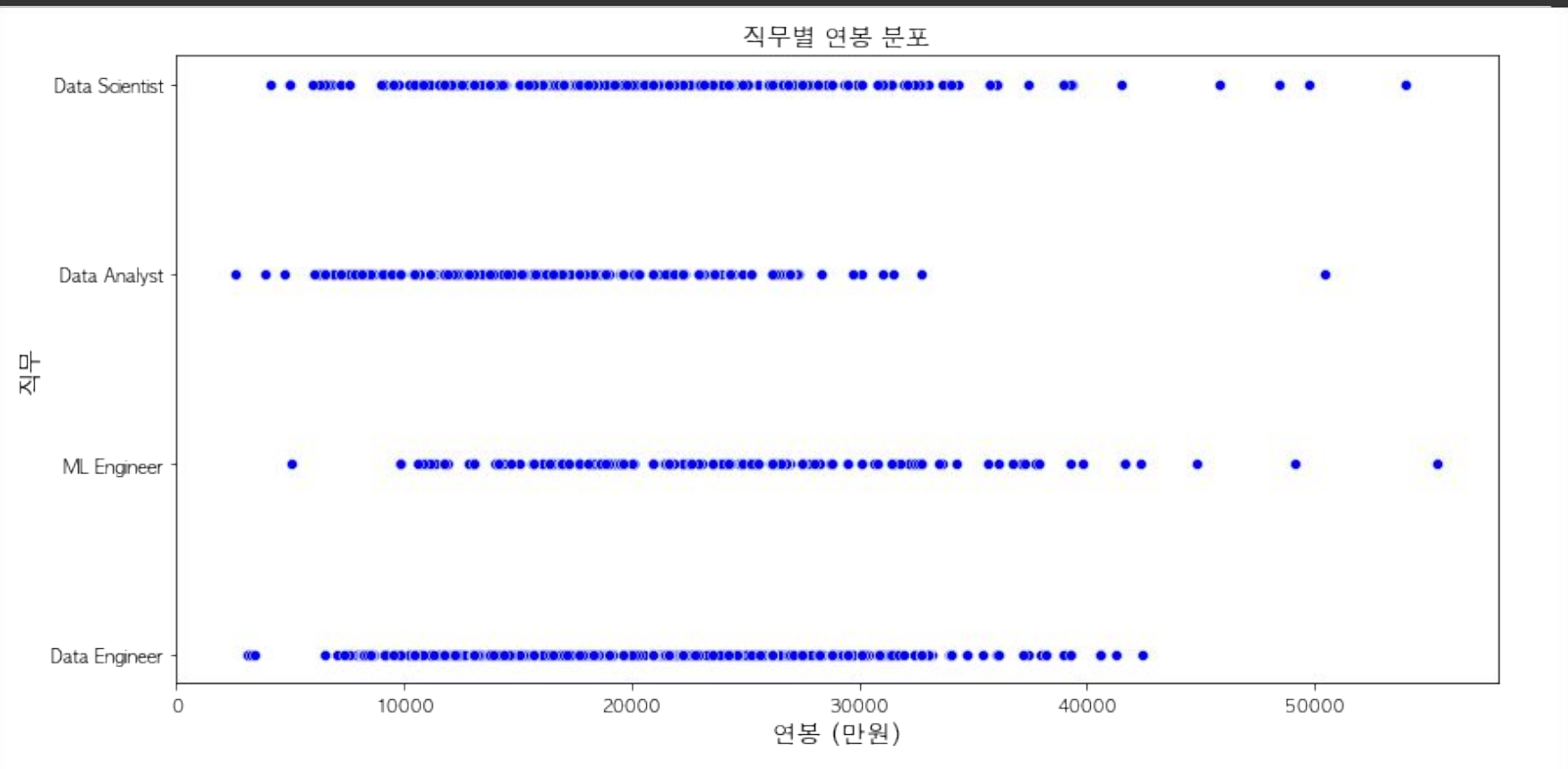


# - 직무별 연봉 분포

직무별 연봉 분포



# — 직무별 연봉 분포





# 결론



감사합니다