

# DATA MULTITASKER

---

**BAK SUNG HO**

---



NLP, MACHINE LEARNING, SCRAPING, RPA, ENGINEERING, DASHBOARD, GROSSH

# 효율과 소통을 중요시하는 데이터 분석가



박성호 (1996.07.29)



010 - 3236 - 6724



datapark1018@gmail.com



신한대학교 글로벌통상경영 (2017~2023)



ADsP, SQLD, TOEIC 870

끊임없이 질문하고 프로세스를 개선하는 DRI 마인드를 소유한 데이터 분석가 입니다.

단순히 그래프와 보고서를 작성하는 것을 넘어서 데이터 분석 프로세스를 점검하고,

반복적인 업무는 자동화하고, 프로세스를 보다 효율적으로 개선하는 것에 강점이 있습니다.

여러 분야의 팀 프로젝트를 빌딩하고 리드하며 쌓은 경험으로, 타 분야의 사람과 협업에 능하고

다양한 BM에 빠르고 유연하게 적응할 수 있습니다.

## CAREER

### 빌트온 (2022.12~2023.03)

- 데이터 품질 관리 및 개선
- 검수 프로세스 통합 및 Python 컨버전
- 이상치 탐지 및 데이터 정제 자동화

### 한국수퍼유통사업협동조합 (2023.07~2024.04)

- 데이터 분석 및 보고서 작성
- 주류 유통 시장 데이터 대시보드 작성
- 데이터 검수 및 분석 프로세스 자동화

# MY CAREER

**Builtón** 글로벌 e-commerce 데이터 분석 솔루션 기업

주 사업내용 : Shopee, Lazada 등 유명 외국 온라인 쇼핑 데이터 솔루션

담당 업무 : 데이터 품질 검수 및 개선 ( DATA QI팀 사원 )

## 주요 성과

- 국가/상품에 따라 데이터 스키마에 따라 독립된 100여개 검수 프로세스 통합
- 통합 META 관리 파일 시스템 생성 및 파이썬 컨버전
- 중복 SKU 검출 프로세스 고도화 (기존 대비 1000% 검출량 확대)

## 중복 SKU 검출 일부

|     |                        |         |           |
|-----|------------------------|---------|-----------|
| 424 | Shenzhen Qianzimei Cos | Lanbena | Lanbena   |
| 424 | Shenzhen LANBENA Cos   | LANBENA | LANBENA   |
| 424 | Shenzhen Sharling Tech | Lana    | Lana      |
| 424 | Lana Bio cosmetics     | LANA    | LANA      |
| 562 | Lactus                 | Lactus  | Lactus Se |
| 403 | The SkinFactory Co.    | KUNDAL  | Kundal    |
| 376 | First Cosmed           | Kiori   | Kiori     |
| 376 | Fine Skin (Thailand)   | Kiori   | Kiori     |
| 452 | Kao Corporation        | Kao     | Kao       |
| 452 | Kao                    | Kao     | Kao       |
| 517 | Kao Corporation        | Kao     | Kao       |

## R to Python

```
if(nrow(pattern_p_b$cu
)) & exists("pattern_p_b$cu
pattern_p_b$cu
pattern_p_b$cu
for(j in 2:nrow(pattern_p_b$cu
pattern_p_b$cu
for k in notebook.tqdm(range(len(categories)
print(k+1, '번째 실행 중 : ',categories)
query = "select seq from mass_md_seq
final_seq = pd.read_sql_query(query,
if len(list(final_seq['seq'].unique()
print('금주 배치 데이터 없음 : ',c
continue
```

**KVC** 국내 가정용 주류 유통사업 협동조합

주 사업내용 : 국내 주류 유통회사의 상품 및 지역별 매출 데이터 솔루션

담당 업무 : 데이터 분석 리포트 작성 및 시스템 관리 ( IT사업팀 사원 )

## 주요 성과

- 분기별 데이터 분석 리포트 자동화 (80page 가량의 분석 보고서 자동화)
- 데이터 ETL 과정 중 비효율 프로세스 개선 IO 시간 기존대비 60배 절감
- 국내 주류 유통 가격 실시간 언론 모니터링 시스템 개발

## 보고서 자동화



## 가격 노출 모니터링

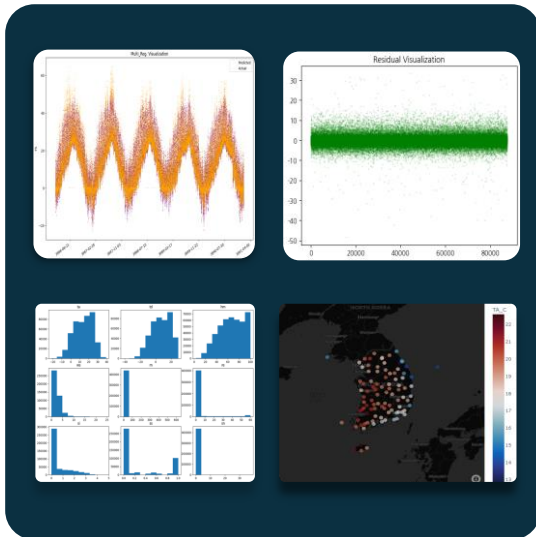
```
time_final = time_final.find_element(By.CSS_SELECTOR, "li > a")
time_final = time_delta(time_final)
driver.execute_script("window.scrollTo(0, document.body.scrollHeight);")
driver.implicitly_wait(3)
time.sleep(1)

temp = BeautifulSoup(driver.page_source, 'html.parser')
temp.publishers = temp.find_all('div', {'class': 'publisher'})
```

| E   | F                  | G     | H   | I     |
|---|--------------------|-------|-----|-------|
| 링크  | 입력일                | 본문    | 출고기 | 가격(원) |
| <a href="http://www.2024-03-08 15:02:4 사진 오비">http://www.2024-03-08 15:02:4 사진 오비</a> | 2024-03-08 15:02:4 | 사진 오비 |     |       |
| <a href="https://n.2024-03-08 14:58:0 쿠팡플레">https://n.2024-03-08 14:58:0 쿠팡플레</a>     | 2024-03-08 14:58:0 | 쿠팡플레  |     |       |
| <a href="http://www.2024-03-08 14:52:4 오비맥주">http://www.2024-03-08 14:52:4 오비맥주</a>   | 2024-03-08 14:52:4 | 오비맥주  |     | 750원  |
| <a href="http://www.2024-03-08 14:50:4 오비맥주">http://www.2024-03-08 14:50:4 오비맥주</a>   | 2024-03-08 14:50:4 | 오비맥주  |     |       |
| <a href="http://www.2024-03-08 14:50:4 오비맥주">http://www.2024-03-08 14:50:4 오비맥주</a>   | 2024-03-08 14:50:4 | 오비맥주  |     |       |

# PROJECTS

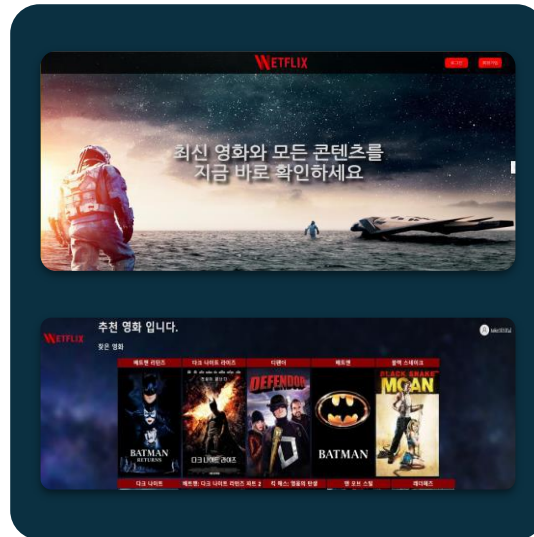
## 지표면 온도 산출 모형 개발



기상청 날씨마루 빅데이터 콘테스트 출품작.  
5개년간의 기상 데이터 (온도, 습도, 풍속 등) 10개를  
활용하여 지면 온도를 산출하는 모형 개발

1. 기상 데이터 수집 (기상청 API, 날씨마루)
2. feature 특성에 따른 시공간 요소 반영, 결측치 보간
3. 데이터 모델링 비교 (LSTM, XGBoost, ...)
4. 분석 결과 확인 및 모델 개선

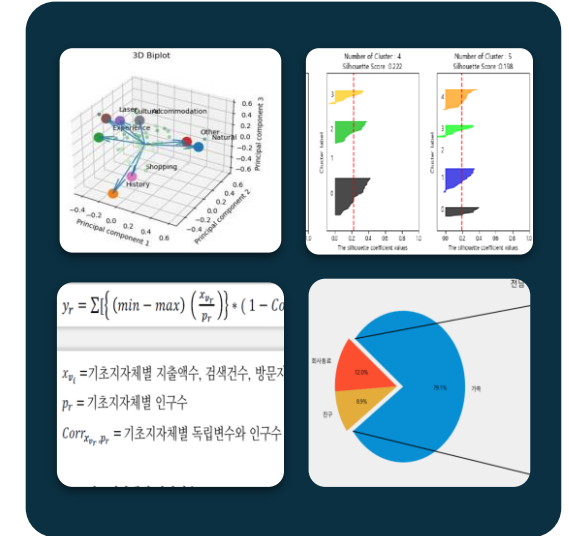
## 영화 추천 웹 서비스 구현



Kaggle의 유명 Competition인 'TMDB 5000' 의  
영화 추천 알고리즘을 Web으로 구현한 토이 프로젝트  
파이썬 웹 프레임 워크 Django 활용.

1. 로컬 서버로 기본 유저 CRUD 구현
2. TMDB API 활용 데이터 수집 및 전처리
3. 영화 줄거리 유사도 분석 (TF-IDF)
4. 사용자 선호도 기반 가중치 부여

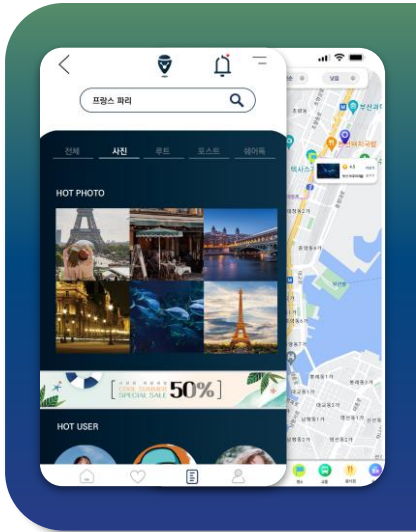
## 관광 활성화 전략 분석



한국관광공사의 데이터 콘테스트 출품작.  
인구소멸 지역의 생활 및 유동인구 증가 전략 분석

1. 시군구별 데이터 수집 (공공데이터, 통계청 등)
2. feature 전처리 및 스케일링
3. 상관관계 분석 및 차원 축소
4. 데이터 모델링 (다중회귀, K-means)
5. 지역별 관광 활성화 전략 도출

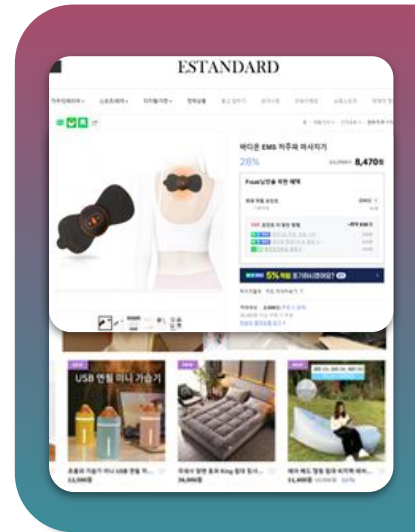
# OTHER EXP



## ✓ 맞춤형 여행 루트 추천 모바일 앱

창업 아이디어 경진대회 출품작.  
사용자 맞춤형 여행 루트 추천 모바일 앱으로서  
User의 정보를 수집하고 이를 기반으로  
동 연령대, 성별, 구성원 등을 구분하여 관광지를 직선으로 이은 코스가 아닌 자세한 GPS 루트를 추천 어플 사업계획

1. 사업계획 구상 및 제안서 작성
2. 경쟁 플랫폼(Triple, 마이리얼트립) 분석
3. 개발자 / 디자이너 팀원 모집 및 구성
4. UI / UX 설계 및 디자인



## ✓ 아이디어 상품 스마트 스토어

아이디어 상품 스마트 스토어 개설.  
e-commerce 관련 교육 수료 후 스마트 스토어 개설.  
사업 제안서 및 아이템 소싱, 상품 상세, 온라인 마케팅 등  
e-commerce 프로세스 구현 및 습득

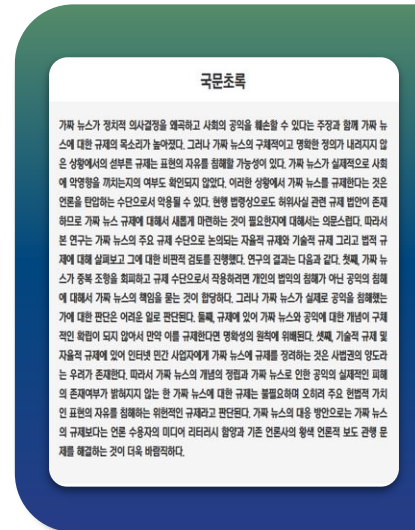
1. 교내 입주기업 선정 및 자금 조달
2. 상품 선정 및 소싱 업체 선정
3. 마케팅 채널 설립 및 상품 상세 제작
4. 스마트 스토어 개설 및 판매



## ✓ 사회 이슈 관련 토론 대회

사회적 이슈가 되는 논제들을 탐구하고자 토론 대회 참여  
팀원 모집, 팀 리더, 토론 전략서 작성.

- 인구보건 토론회 본선  
⇒ 국내외 출산 장려 정책, 시민 인식 조사, 청년 코호트 분석
- 계명 토론회 본선  
⇒ 통일 관련 경제적 비용 및 효과 분석
- 통상정책 토론회 본선  
⇒ 미중 무역 수치 및 보호주의에 따른 무역장벽 분석



## ✓ 가짜 뉴스 논문 공모전

가짜 뉴스 관련 논문 공모전 출품작  
소셜 플랫폼의 성장과 stable-diffusion와 같은 딥 러닝 모델을 사용한 deep fake의 등장에 따른 가짜 뉴스 확산의 문제를 규제하는 방법에 대한 탐구 보고서.  
법적 규제의 경우 피해 주체 특정 모호성과 언론자유권의 권리 충돌의 문제점에 대해 연구. 자율적 규제의 경우, ISP 사업자의 알고리즘 또는 수동적 규제에 대한 이용자의 권리침해 등을 연구



# Thank you for reading

## If you're interested, Please contact me



010 – 3236 – 6724



[datapark1018@gmail.com](mailto:datapark1018@gmail.com)