

PaaS 환경에서 다중 사용자 머신 러닝 플랫폼
구축 데모

- 목적

- 데이터 분석이 필요한 사용자에게 수십초내에 동일한 컨테이너 기반의 분석 환경을 제공하는 것을 보여줍니다.

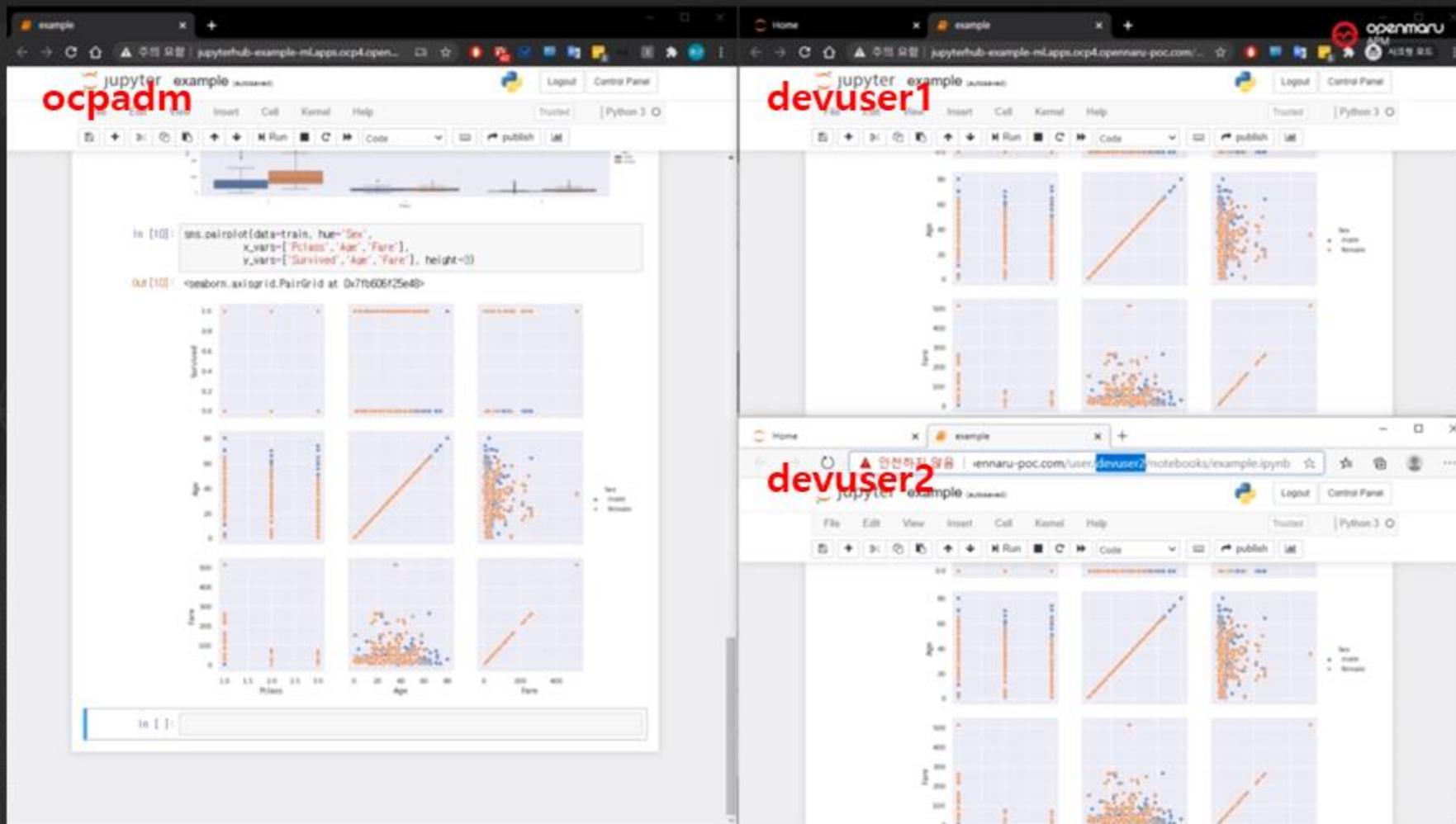
- 내용

- 이 데모에서는 머신러닝 개발환경으로 가장 많이 사용하는 주피터노트북을 사용자에게 제공하고 사용자는 타이타닉 승선 기록 데이터를 가지고 분석/가공하여 데이터를 원하는 형태의 테이블과 차트로 만드는 것을 보여줍니다.

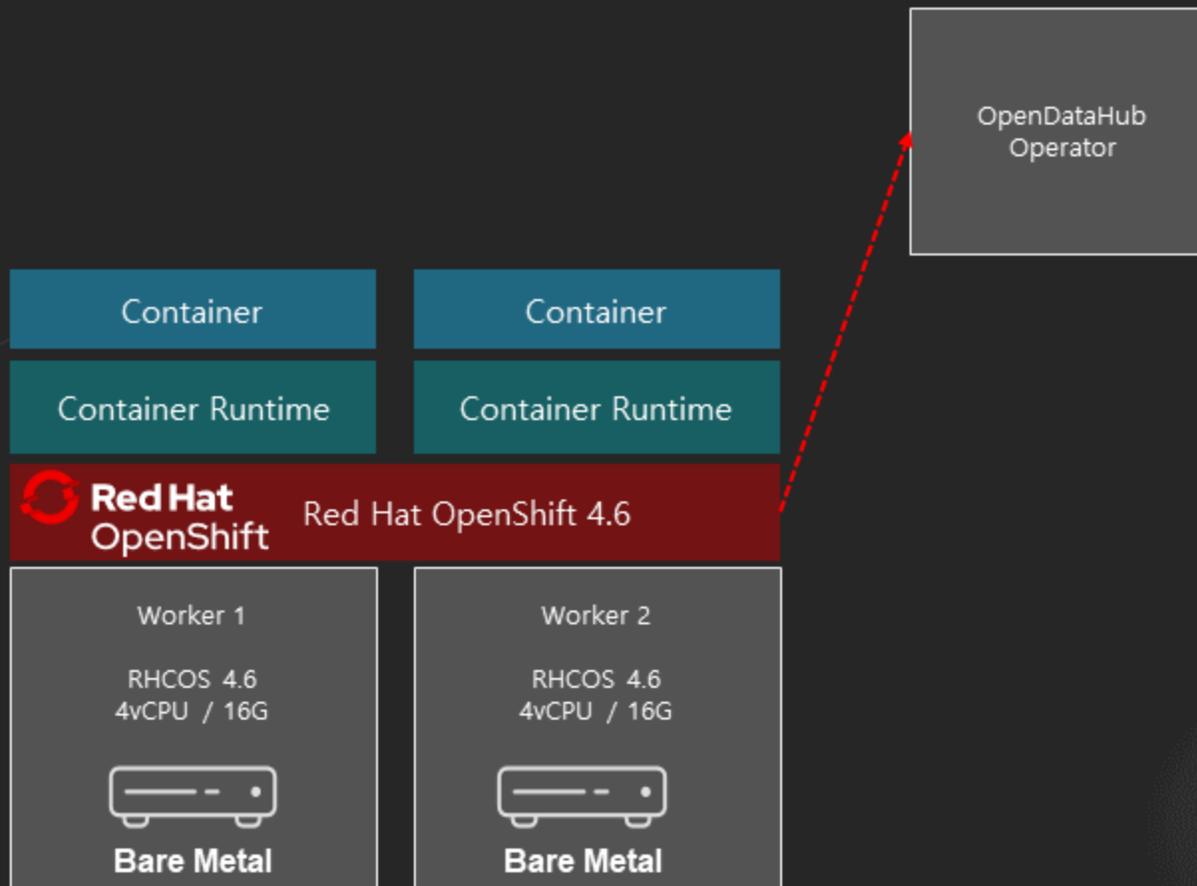


머신러닝 개발 환경을 VM으로 관리하기 어려워요.
PaaS 환경에서 머신러닝 개발 환경을 쉽게 제공할 수 있는 방법은 없을까요?

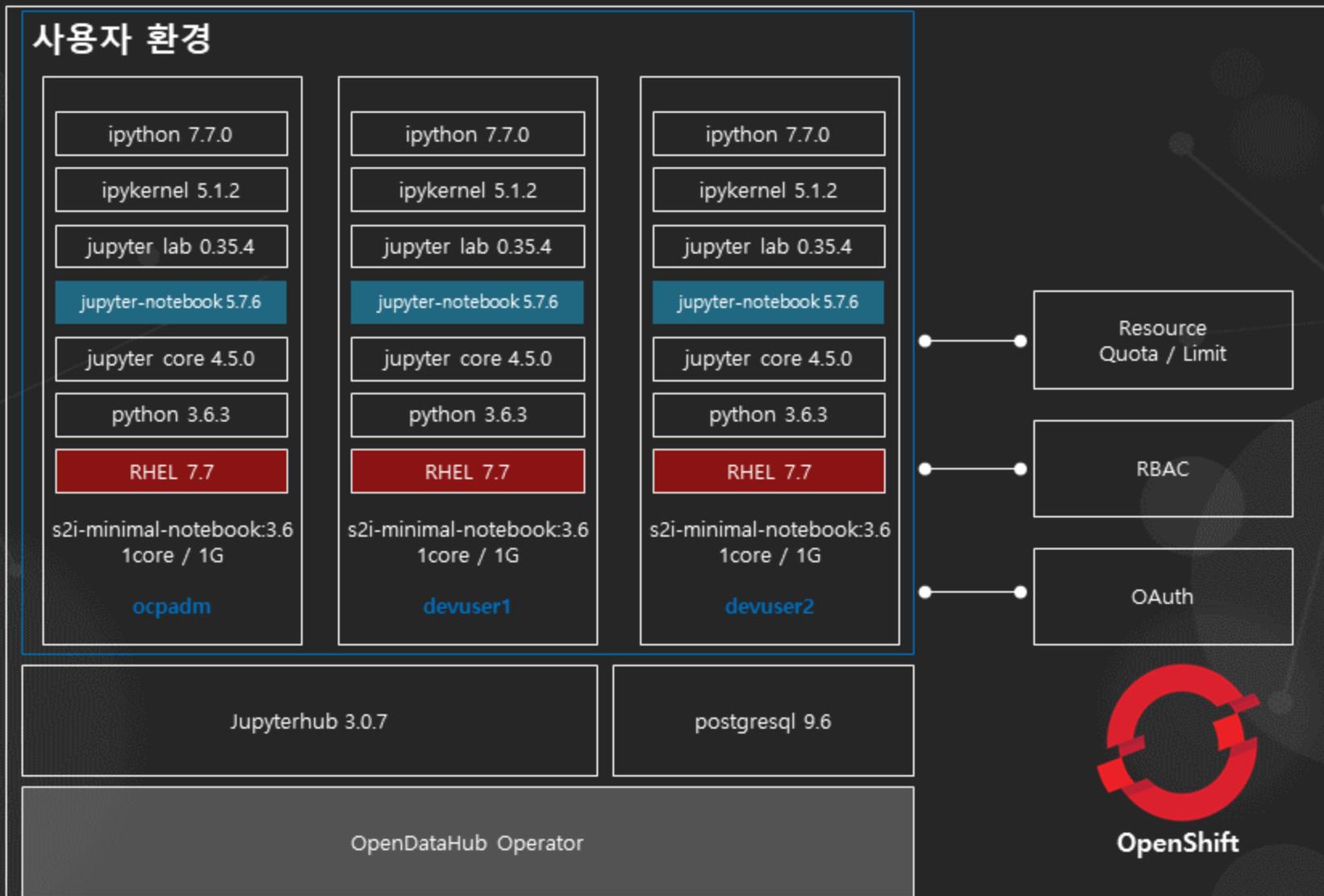
PaaS 환경에서 다중사용자에게 머신러닝 환경 제공



Container 환경 테스트 구성도



주피터 SaaS Service 구성도



신규 서비스 구축

PaaS 환경에서 다중 사용자
머신 러닝 플랫폼 구축



OpenDataHub Operator 설치



The screenshot shows the OpenShift console interface for the 'example-ml' project. The left sidebar contains navigation menus for Administrator, Home, Projects, Operators, Workloads, and more. The main content area displays the 'Project Details' for 'example-ml', which is in an 'Active' state. The 'Overview' tab is selected, showing a summary of the project's status and resource utilization.

Project Details:

- Name: example-ml
- Requester: ocpadm
- Labels: No labels

Status: Active

Utilization (1 Hour):

Resource	Usage	16:30	16:45	17:00	17:15
CPU	180.3m	500m	~500m	~500m	~500m
Memory	4.16 GiB	6 GiB	~4 GiB	~4 GiB	~4 GiB
Filesystem	43.15 MiB	60 MiB	~40 MiB	~40 MiB	~40 MiB
Network Transfer	0 Bps in, 0 Bps out	20 KBps	~10 KBps	~10 KBps	~10 KBps
Pod count	1	15	~10	~10	~10

Inventory:

- 0 Deployments
- 0 Deployment Configs
- 0 Stateful Sets
- 0 Pods
- 0 PVCs
- 0 Services
- 0 Routes
- 0 Config Maps
- 12 Secrets

Activity: No ongoing activities or recent events.

OpenDataHub Operator 설치.

신규 서비스 구축



머신러닝 개발환경 활용

train.csv

```
승객아이디,생존여부,좌석등급,이름,성별,나이,형제 자매 / 배우자 수,부모와 자식 수,티켓번호,요금,객실 번호,정박 항구  
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,7.25,,S  
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thayer)",female,38,1,0,PC 17599,71.2833,C85,C  
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2. 3101282,7.925,,S  
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",female,35,1,0,113803,53.1,C123,S  
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8.05,,S  
6,0,3,"Moran, Mr. James",male,,0,0,330877,8.4583,,Q  
.  
.  
891,0,3,"Dooley, Mr. Patrick",male,32,0,0,370376,7.75,,Q
```

테스트데이터 출처 : kaggle
<https://www.kaggle.com/c/titanic/data>

example.ipynb

```
import pandas as pd  
train = pd.read_csv('train.csv')  
train.head()  
  
import matplotlib.pyplot as plt  
import seaborn as sns  
sns.set()  
  
sns.relplot(data=train, x='Age', y='Survived')  
  
sns.relplot(data=train, x='Age', y='Survived', hue='Sex', aspect=3).  
.  
.  
.
```

샘플코드 수행



Log in to your account

Red Hat
OpenShift Container Platform

Welcome to Red Hat OpenShift Container Platform.

Username *

Password *

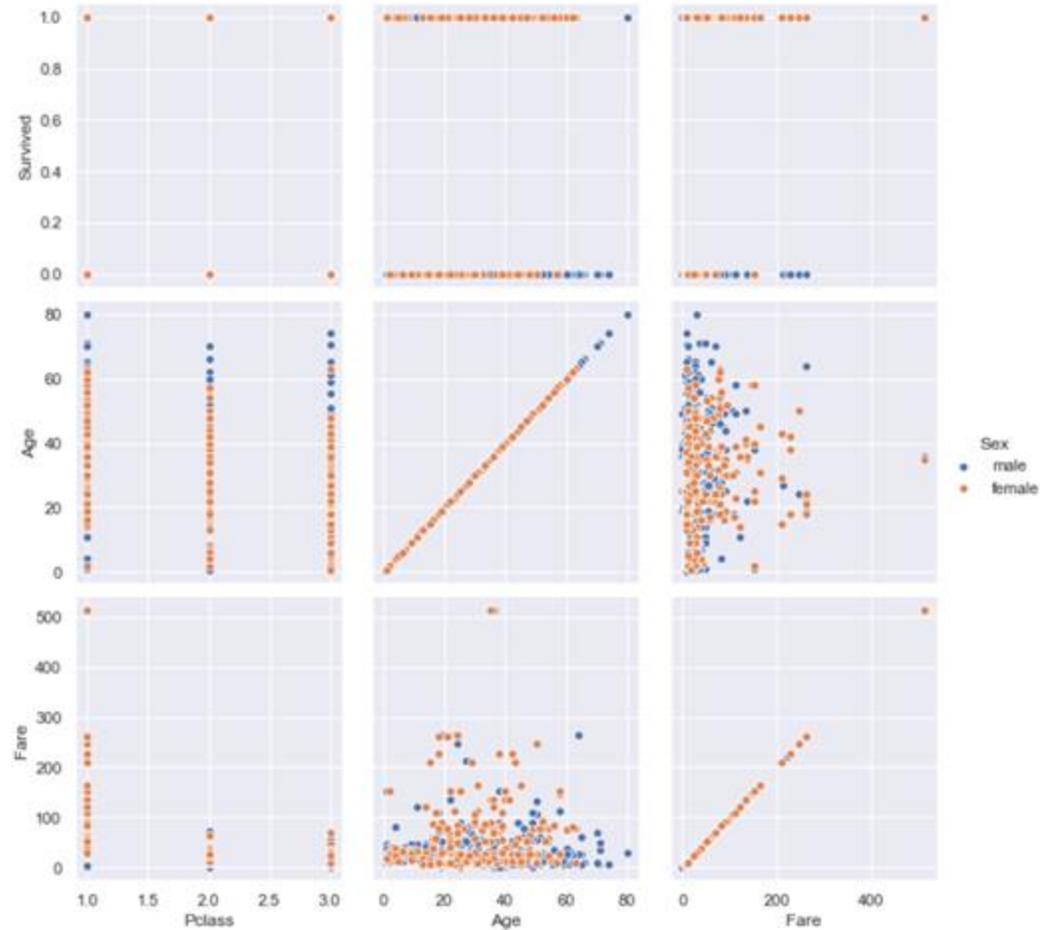
Log in

ocpadm 계정으로 샘플코드 수행.

수행결과 예시

```
In [10]: sns.pairplot(data=train, hue='Sex',  
                    x_vars=['Pclass', 'Age', 'Fare'],  
                    y_vars=['Survived', 'Age', 'Fare'], height=3)
```

Out [10]: <seaborn.axisgrid.PairGrid at 0x2a848a90580>





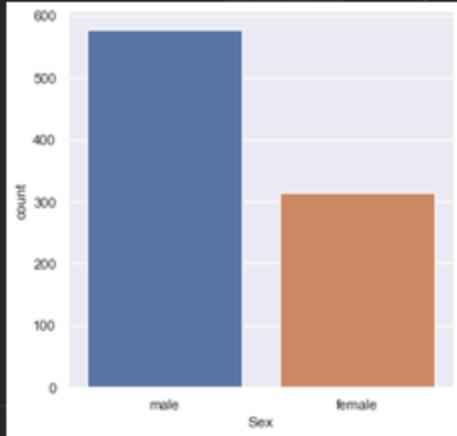
신규 서비스 구축

머신러닝 개발환경을 통한
데이터분석 결과

데이터 분석 결과

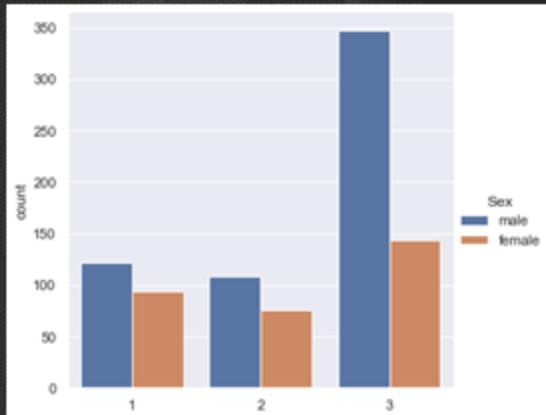
탑승자 성비

- `sns.factorplot('Sex', kind='count', data=train)`
- 탑승인원 성비는 남성,여성 2:1



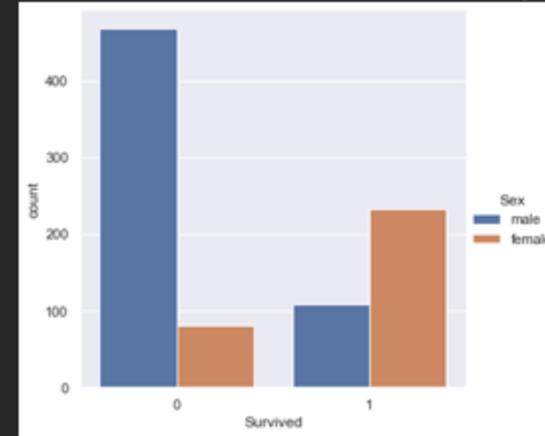
좌석등급별 성비(낮을수록 높은 등급)

- `sns.factorplot('Pclass', kind='count', hue='Sex', data=train)`
- 최하위 등급 좌석의 승객은 남성이 2배 이상



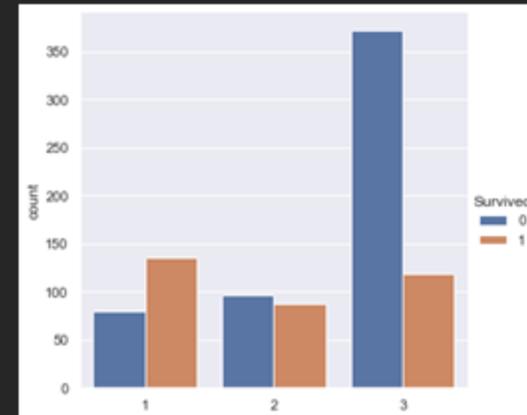
성별 생존률(1이 생존)

- `sns.factorplot('Survived', kind='count', hue='Sex', data=train)`
- 생존확률은 남성이 약 20%, 여성은 약 60%



좌석등급별 생존률(1이 생존)

- `sns.factorplot('Pclass', kind='count', hue='Survived', data=train)`
- 최상위 등급 좌석은 생존률 60%, 최하위 등급 좌석은 생존률 20%





openmaru

제품 / 서비스에 관한 문의

- 콜 센터 : 02-469-5426 (휴대폰 : 010-2243-3394)
- 전자 메일 : sales@openmaru.com