



한적하게 벚꽃을 보려면 언제 연차를 내는 것이 좋을까?

☼ 상태	완료
≡ 월	4월
≡ 문제 상황	한적하게 벚꽃을 보러가고 싶은데 사람이 너무 많으니까 한적하게 벚꽃을 볼 수가 없음. 그래서 사람이 적을 때 벚꽃을 풍성하게 보고 싶음.
≡ 해결하고자 하는 목적	비가 오지 않는 시기 + 한적한 시기를 도출해서 언제쯤 휴가/휴일을 신청해야 벚꽃을 보러갈 수 있을 지 시기를 제안한다.
≡ 필요 데이터	10년 간 기상청 강수량, 풍량 데이터, 10년 간 4월 석촌호수 일대 벚꽃 개화, 지는 시기 수집, 10년 간 석촌 호수 일대 인구 데이터

분석 배경

- 최근 벚꽃 보러 갔는데, 사람이 너무 많았어서 힘들었음
- 나는 한적한 날에 보고 싶은데, 마침 여자친구는 평일에도 휴일 신청을 할 수 있음
- 내년에는 여자친구한테 벚꽃이 필 때 좀 한적한 시기를 과거의 데이터를 통해 도출해서 미리 연차를 낼 수 있는 근거를 마련하여 우리끼리 한적하게 벚꽃을 보고자 함

분석 목표

- 언제 벚꽃을 가장 예쁘고 편하게 볼 수 있는 지 도출하여 연차를 낼 시기를 예측할 수 있도록 한다.

예쁘게 본다는 정의

- 날씨가 좋은 날 (벚꽃이 떨어지지 않는 강풍이 불지 않는 날, 비가 안오는 날)
- 사람이 다른 날보다 비교적 적은 날
- 개화일 당일이 아닌 날

분석 방법

1. 데이터 셋은 서울시 열린 데이터 광장에서 2021~2024년도 간 스마트 도시 데이터 센터 1시간 당 유동인구 를 측정한 데이터를 활용했고, 석촌호수 지역에 해당하는 코드만 필터링해서 가져왔다.
2. 기상청에 저장된 서울시의 10분 단위로 저장된 강수량, 풍속 데이터를 가져왔다. 그리고 두 데이터 전처리 하여 날짜 기준으로 병합하여 하나의 데이터 프레임을 만들었다.
3. 4년 간의 잠실 석촌호수 일대의 벚꽃 개화 시기, 지는 시기를 수집하여 데이터 프레임 칼럼에 라벨 형태로 추가하였다.
4. 비 오는 날(강수량이 0 이상), 강풍이 부는 날(6m/s 이상 일 경우)은 벚꽃이 떨어지므로 이 날 이후의 데이터들은 제거하였다.
5. 날씨와 요일 간의 상관 관계, 인과 관계가 있는 지 간단히 살펴본다.
6. 인과 관계가 유효하다면, 시기 별 방문자 수 확인, 요일 별 평균 방문자 수를 태블로로 시각화 하여 인사이트를 도출한다.

시각화 방법

1. (연도 별) 벚꽃 시기 별 방문자 수 확인
2. (요일 별 + 시간대 별) 평균 방문자 수 확인

연차를 낸다고 가정할 때 벚꽃을 가장 보기 좋은 시기는 언제일까?

🎯 알아둬야 할 점

우선 알아둬야 할 것이 있다. 벚꽃이 개화 한 후, 5~7일 이후 절정에 이르고, 절정에 이른 후, 3~5일 뒤에 벚꽃이 낙화하게 된다. 데이터 분석에 필요한 벚꽃의 개화, 절정, 낙화시기는 다음과 같고, 케이웨더의 벚꽃 개화시기를 참고했다.

케이웨더

 https://kweather.co.kr/sub/intro/press_board.php?id=143



2025년 4월 4일 개화, 4월 9~11일 절정, 4월 14~16일 벚꽃 짐

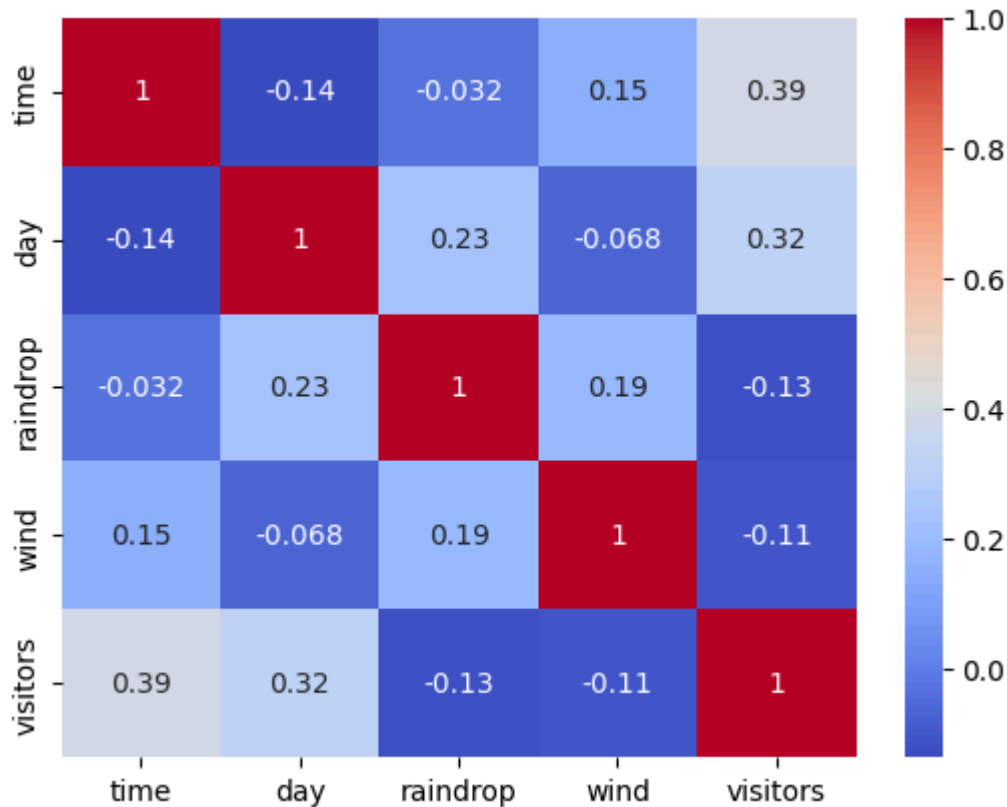
2024년 4월 1일 개화, 4월 6~8일 절정, 4월 11~13일 벚꽃 짐

2023년 3월 25일 개화, 4월 1~3일 절정, 4월 6~8일 벚꽃 짐
2022년 4월 4일 개화, 4월 9~11일 절정, 4월 14~16일 벚꽃 짐
2021년 3월 24일 개화, 3월 29~31일 절정, 4월 3~5일 벚꽃 짐

따라서, 이 시기에 해당하는 데이터들만 가져왔다. 그리고, 벚꽃은 풍속 6m/s 이상일 때, 그리고 강수량이 10mm 이상일 경우 꽃이 진다. 따라서, 이 경우에 해당하는 데이터인 2021년 4월 3일 이후의 데이터(일 강수량 40mm, 바람 6m/s이상), 2023년 4월 4일 이후의 데이터(일 강수량 26.1mm)는 삭제했다.

분석 내용

상관 관계 분석



1. time(벚꽃 개화 시기), day(요일)에 따라 상관 계수가 각각 0.39, 0.32으로 양의 상관 관계를 보인다.
2. 벚꽃이 더 피어날수록, 사람도 더 많이 늘어나고 주말에 다가갈수록 석촌 호수의 방문자 수는 늘어가는 선형적 관계가 있다는 뜻이다.

인과 관계 분석

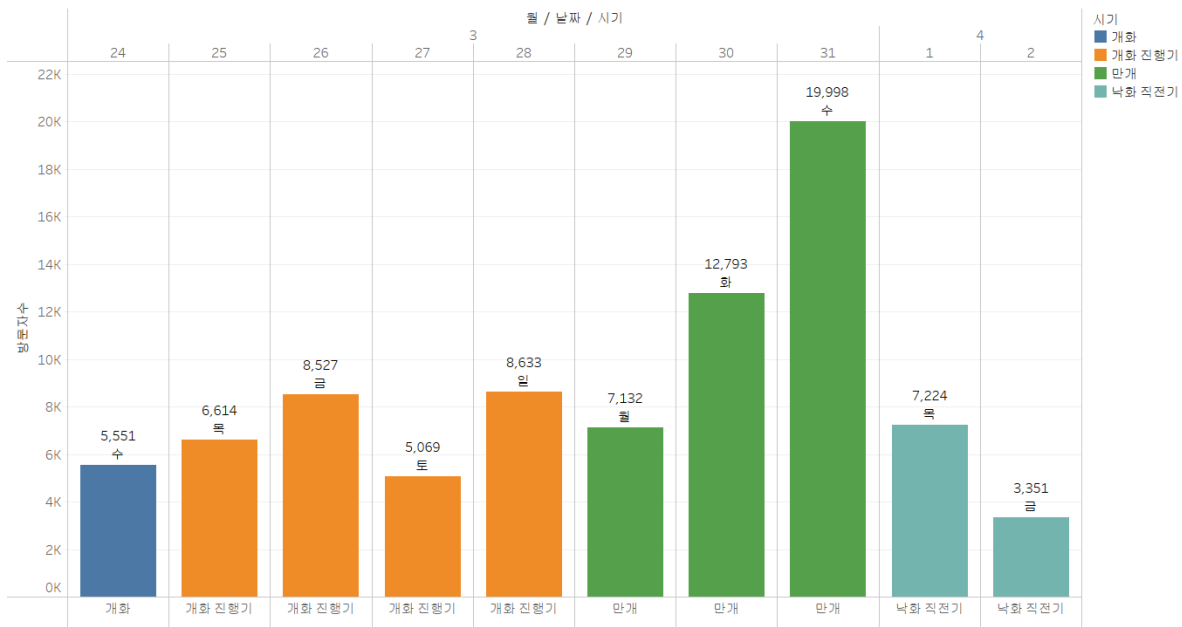
	coef	std err	t	P> t	[0.025	0.975]
const	-1160.0439	6432.990	-0.180	0.858	-1.42e+04	1.18e+04
time	3367.3831	939.751	3.583	0.001	1469.518	5265.249
day	1964.7232	623.677	3.150	0.003	705.182	3224.264
raindrop	-359.2276	260.031	-1.381	0.175	-884.370	165.915
wind	-1733.0810	2008.562	-0.863	0.393	-5789.455	2323.293

1. 벚꽃 개화 시기는 단계가 1단계 증가(예를 들면, 개화 → 개화 진행기, 개화 진행기 → 만개)할수록, 약 3,367명 증가하고, 요일 수가 증가(예를 들면, 월 → 화, 금 → 토)할수록, 약 1,965명이 증가하고 유의하다. ($p < 0.05$)
2. 비나 풍속이 증가함에 따라 방문자 수는 감소, 이 두 계수들은 p-value가 높아 유의하지 않았다!

즉, 벚꽃이 필수록 사람은 많아지고, 주말에 가까워질수록 사람은 많아진다는 것이 증명!

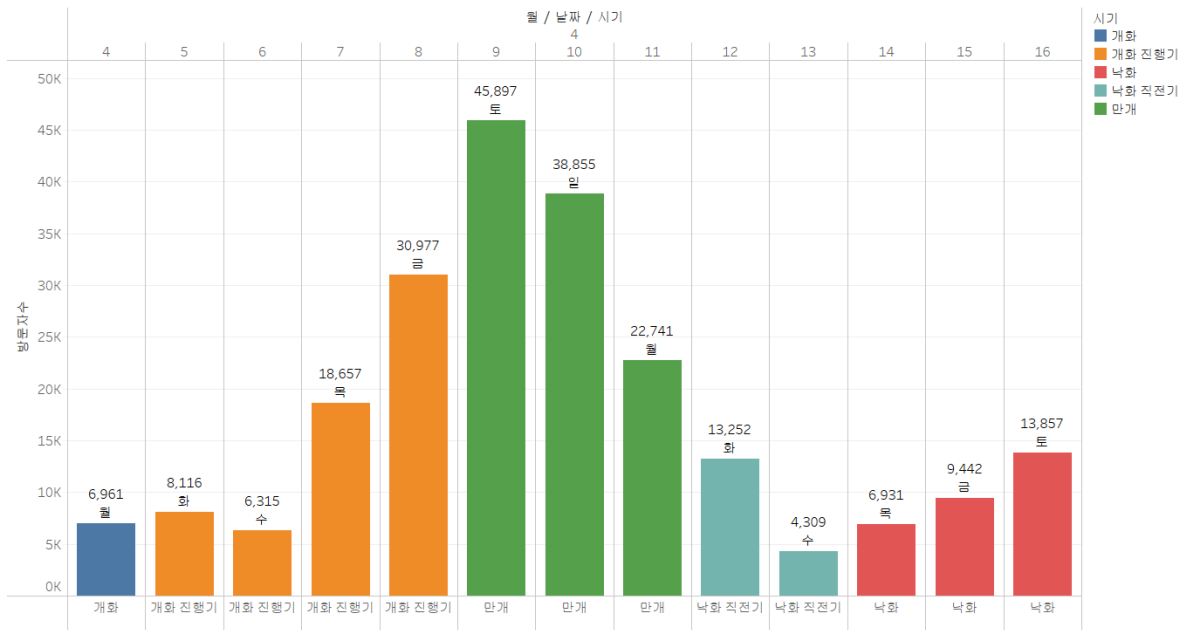
✈️ 그럼 언제 가야할까??

2021년 벚꽃 시기 별 하루 방문자 수



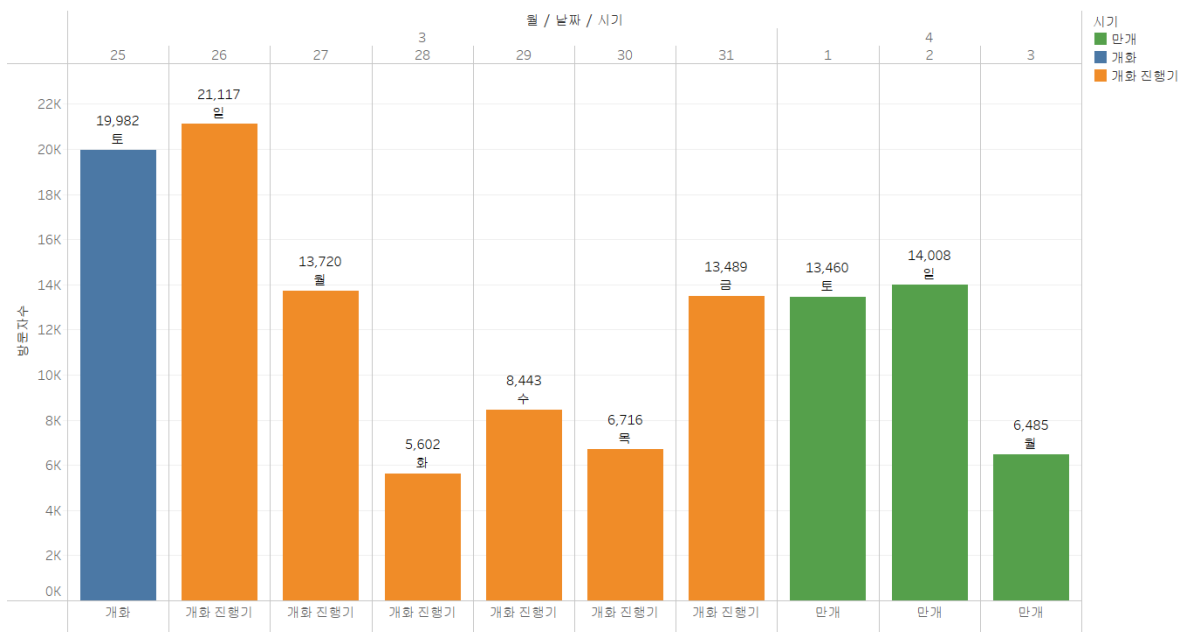
월 및 날짜별로 구분된 각 시기에 대해 방문자수의 합계입니다. 색상이 시기의 세부 정보를 표시합니다. 마크에 레이블이 방문자수의 합계 및 요일에 의해 지정되었습니다. 요일의 세부 정보가 표시되었습니다. 2021년(월) 유지하는 연도에 대한 데이터가 필터링되었습니다.

2022년 벚꽃 시기 별 하루 방문자 수



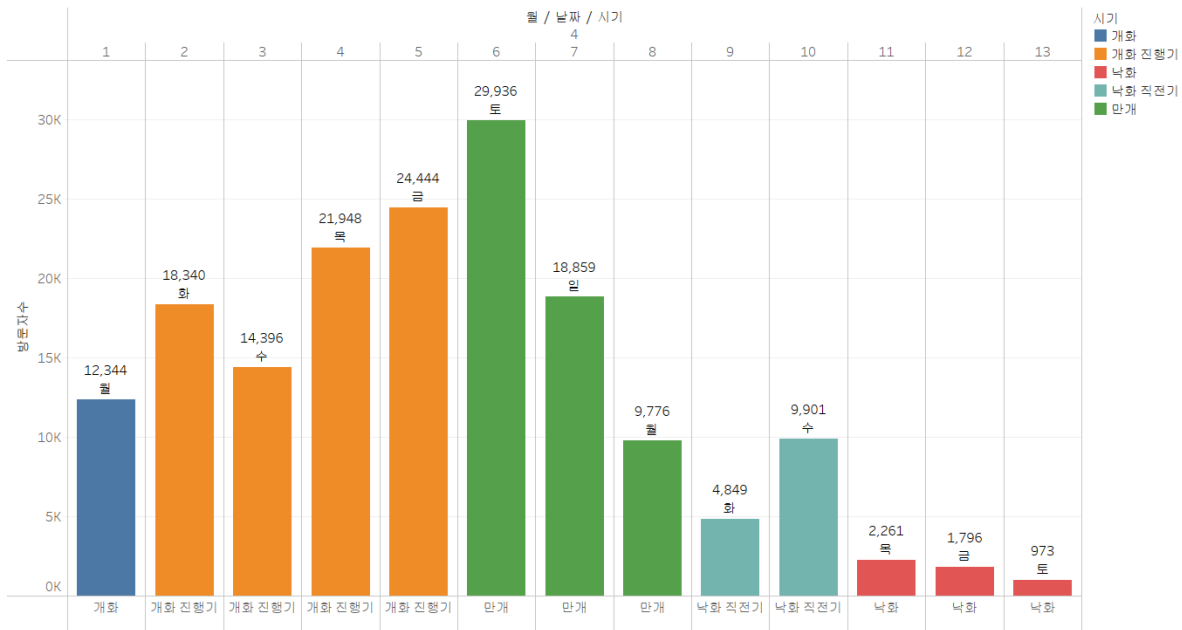
월 및 날짜일별로 구분된 각 시기에 대해 방문자수의 합계입니다. 색상이 시기의 세부 정보를 표시합니다. 마크에 레이블이 방문자수의 합계 및 요일에 의해 지정되었습니다. 2022월(를) 유지하는 연도에 대한 데이터가 필터링되었습니다.

2023년 벚꽃 시기 별 하루 방문자 수



월 및 날짜일별로 구분된 각 시기에 대해 방문자수의 합계입니다. 색상이 시기의 세부 정보를 표시합니다. 마크에 레이블이 방문자수의 합계 및 요일에 의해 지정되었습니다. 요일의 세부 정보가 표시되었습니다. 2023월(를) 유지하는 연도에 대한 데이터가 필터링되었습니다.

2024년 벚꽃 시기 별 하루 방문자 수

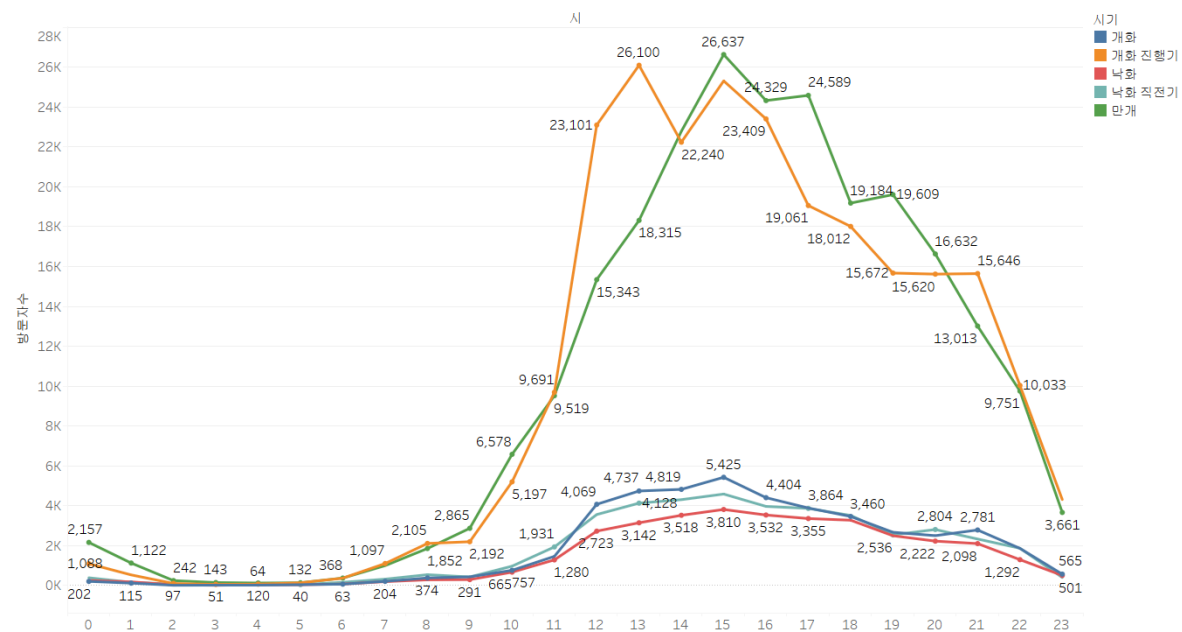


월 및 날짜별로 구분된 시기에 대해 방문자수의 합계입니다. 색상이 시기의 세부 정보를 표시합니다. 마크에 레이블이 방문자수의 합계 및 요일에 의해 지정되었습니다. 요일의 세부 정보가 표시되었습니다. 2024년(월) 유지하는 연도에 대한 데이터가 필터링되었습니다.

- 2023년을 제외(2023년도는 27일에 비가 와서 만개 전보다 개화 진행기에 방문자가 많은 듯하다.)하면 모든 년도가 공통적으로 만개 시기가 되면 전반적으로 방문자 수가 증가
⇒ 따라서, 벚꽃을 한산하게 보고 싶다면 만개 시기는 피하는 게 좋다.

시간대 별 방문자 수

벚꽃 시기 및 시간 별 방문자 수

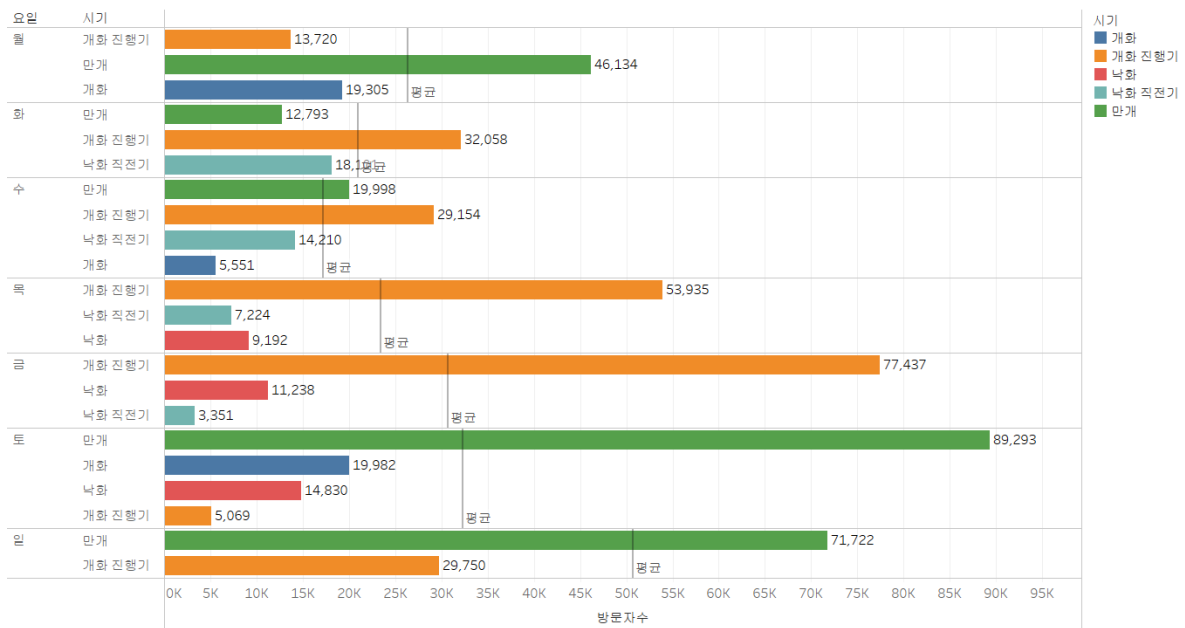


시에 대한 방문자수의 합계의 추세입니다. 색상이 시기의 세부 정보를 표시합니다. 마크에 레이블이 방문자수의 합계에 의해 지정되었습니다. 시기의 세부 정보가 표시되었습니다.

- 전반적으로 오전 10 - 15시 사이 꾸준히 증가하는 성향을 보임.
- 하지만 16시 이후로는 감소하는 추세
- 개화 시기, 낙화 직전기, 낙화 시기는 방문자 수가 대체로 적다.

요일 및 시기 별 방문자 수

요일 및 시기 별 방문자 수



요일별로 구분된 각 시기에 대해 방문자수의 합계입니다. 색상이 시기의 세부 정보를 표시합니다. 마크에 레이블이 방문자수의 합계에 의해 지정되었습니다.

- 토요일은 평균 3만 2천 명, 일요일은 평균 5만 명에 육박 ⇒ 주말인데 벚꽃이 절정이다? 동향차
- 평균적으로 한산한 요일 ⇒ 수요일
- 만개, 개화 진행기에 사람들이 석촌 호수에 사람들이 많이 몰림 ⇒ 비교적 사람이 적은 개화 시기, 혹은 낙화 직전기에 방문하는 것을 추천

최종 제안

진짜 한산하게 벚꽃을 보려고 연차를 쓰고 싶다면?

⇒ 석촌호수 기준, 평균적으로 사람이 적은 수요일에 방문하는 것을 추천

⇒ 시간대는 9시부터 10시 사이 혹은 16시 이후

⇒ 낙화 직전기(만개 시기가 끝난 뒤 1~3일 사이), 혹은 개화 시기에 방문할 것을 권장

만개 시기 일 때 벚꽃이 보고 싶어서 연차를 쓰고 싶다면?

⇒ 무조건 평일에 연차, 주말 연차 X (월요일 추천)

전처리 과정 파일

벚꽃을_잘볼_수_있는_시기를_찾아라.ipynb

데이터 완성본

벚꽃을_잘볼수있는_시기를_찾아라.csv