

4일차 과제

[Day4]

도시대기 미세먼지 데이터와 빅맥 지수 데이터를 이용하여 다양한 시각화를 이용해 “시간 분석”을 해봅시다.

보통 시간 관련된 분석을 할 때는 트렌드를 살펴보는데 유용한 라인차트를 자주 사용하는데요.

라인차트 이외에도 시간과 관련된 데이터를 분석하는 시각화는 다양합니다. 아래 과제를 통해서 한 번 살펴볼까요?

아래 워크북을 다운로드 받으세요.(최신 버전 다운로드를 권장합니다.)

V2020.1 이상 : @DAY4 Time Analysis_Start.twbx

V2018.1 : @DAY4 Time Analysis_Start_v2018.1.twbx

4일차 과제 내용

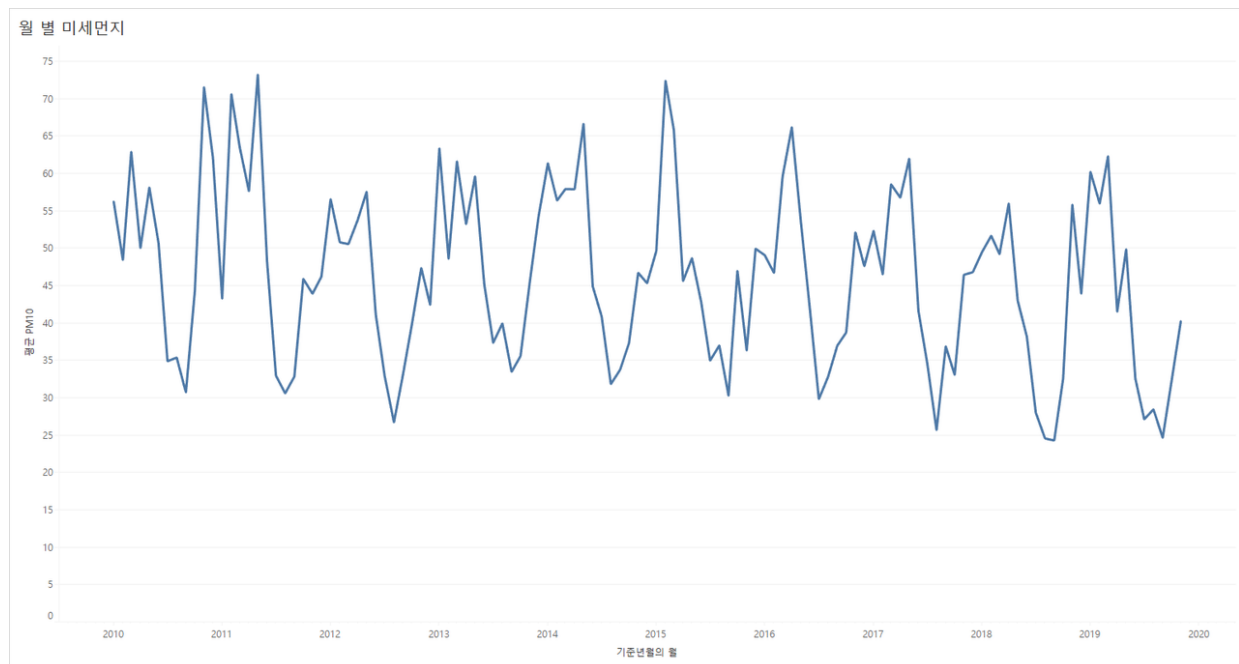
- 1. 라인 차트
- 2. 하이라이트 테이블
- 3. 테이블 계산식을 이용한 시간 분석
- 4. 덤벨차트

1. 라인 차트

라인 차트는 시간의 흐름에 따라 데이터의 패턴과 트렌드를 보는데 가장 많이 사용되는 시각화 입니다.

라인차트를 이용해서 아래와 같이 월 별 미세먼지 평균 지수를 한 번 살펴볼까요 ?

사용 Data : 도시대기 미세먼지(PM10)

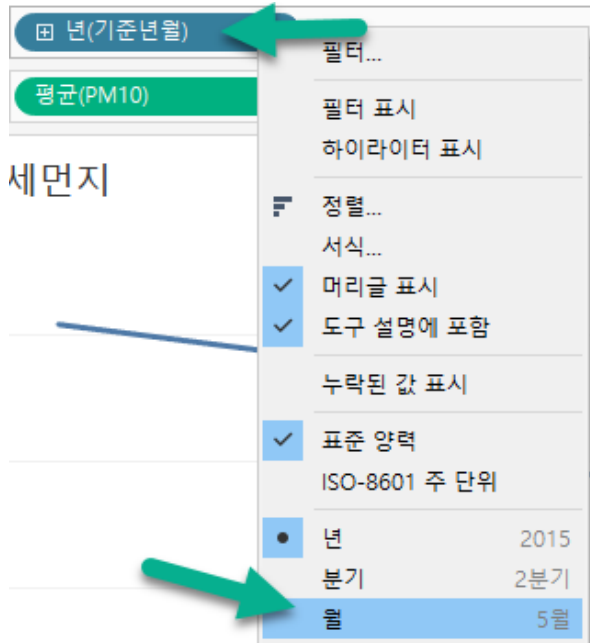


[시각화를 위한 HINT]

1. 기준년월을 열 선반에 놓아주세요.

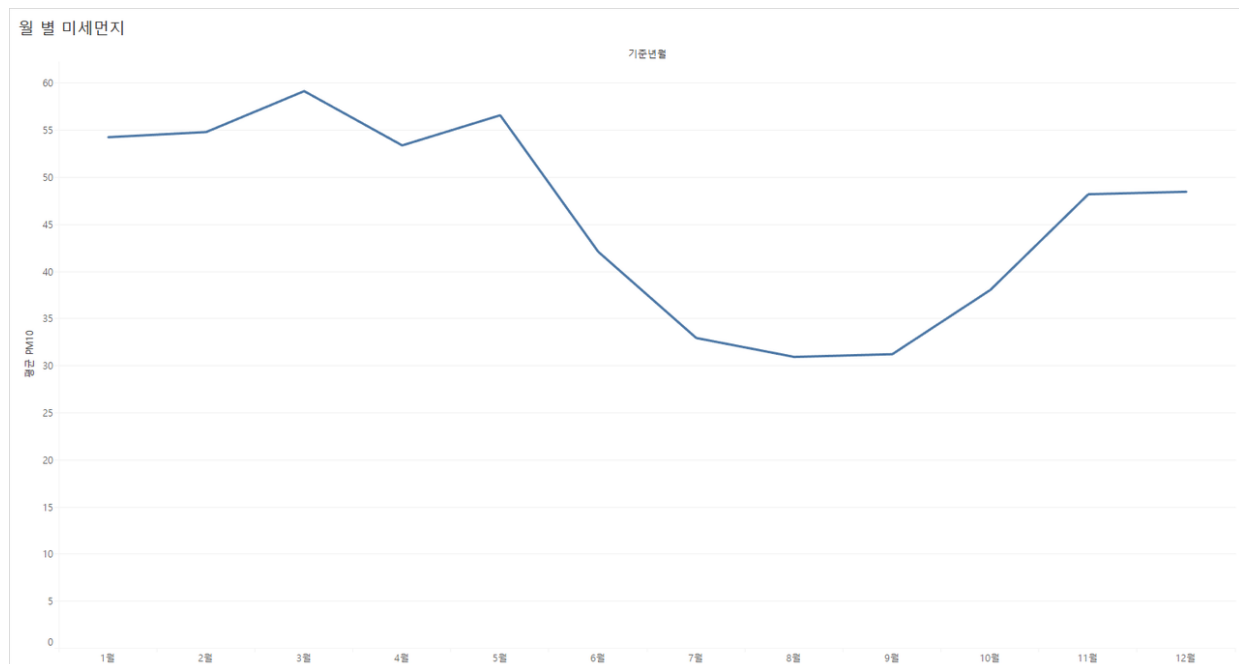
태블로에서는 날짜 필드를 시각화에 가져다 놓으면 기본적으로 “년”을 기준으로 보여줍니다.

이를 “월”로 변경해주세요. 가져다 놓은 기준년월에서 마우스 오른쪽 버튼을 클릭하고 “월”을 선택해주시면 됩니다



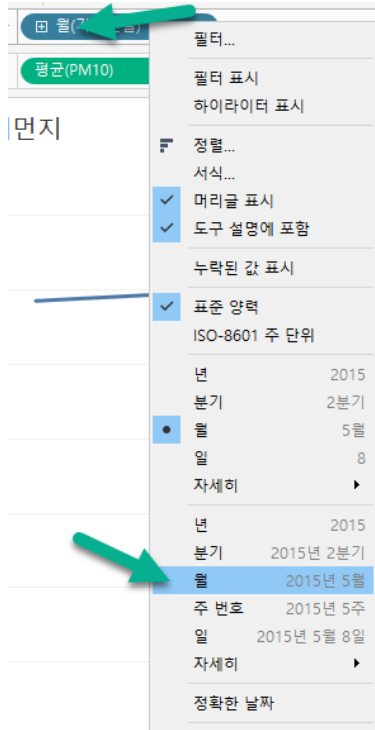
2. PM10의 집계 형태는 평균으로 변경해주세요.

아래 그림처럼, 2010년부터 2019년까지의 데이터를 가지고 각 월 별로 평균 PM10이 나타내진 것을 알 수가 있습니다. 확실히 겨울철에 미세먼지가 심한 것을 확인할 수가 있네요.



3. 위의 시각화처럼 월로 평균 값을 보는 것이 아니라, 2010년도부터 2019년까지의 월 별로 미세먼지가 어떻게 변화했는지 추이를 살펴보고 싶다면 어떻게 해야 할까요?

이번에는 기준년월 알약에서 마우스 오른쪽 버튼을 누르고, 아래 쪽에 있는 월을 선택해보세요.



그럼 이전과는 다르게 시간이 월을 기준으로 연속적으로 표시되는 것을 볼 수 있습니다.

이처럼 날짜의 옵션을 위쪽(불연속형)을 선택하느냐 아래쪽(연속형)을 선택하느냐에 따라서 시간 표현을 다르게 할 수 있습니다. 이 내용은 2주차에 좀 더 자세히 다루니 이런 느낌이구나 정도로 느끼고 지나가시면 좋을 것 같아요~

다음으로 시도 별로 월 별 미세먼지 변화량을 한 번 살펴볼까요? 어떻게 시각화를 해보시겠어요?
한 번 자유롭게 고민해 보시고 다음 과제로 넘어 가볼까요?

2. 하이라이트 테이블

하이라이트 테이블을 이용하여 텍스트에 색상을 추가로 표현하면 여러 값의 패턴을 좀 더 명확하게 볼 수 있습니다.
시도 별 월 별 미세먼지 패턴을 좀 더 명확하게 살펴봅니다.

사용 Data : 도시대기 미세먼지(PM10)

시도별/월별 미세먼지

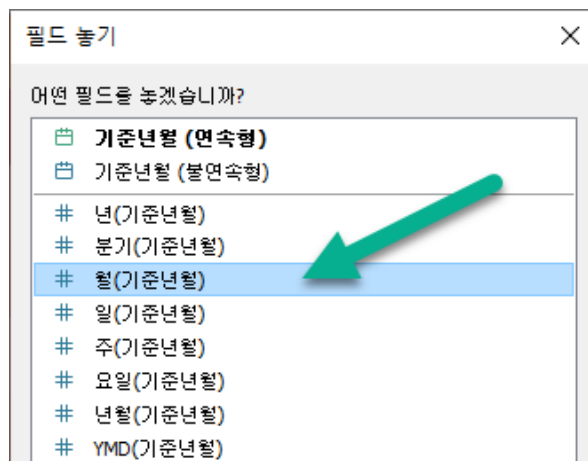
시도	1월	2월	3월	4월	5월	기준년월		6월	7월	8월	9월	10월	11월	12월
강원	58.78	57.64	62.46	53.69	55.01			38.56	30.73	28.98	26.89	33.72	44.59	51.65
경기	62.75	62.42	67.13	59.59	61.74			45.48	35.49	33.27	33.94	41.98	54.04	56.15
경남	48.02	49.65	52.41	49.98	54.12			42.86	35.08	32.90	31.53	37.39	44.96	44.28
경북	51.61	52.23	55.26	51.33	53.27			39.90	32.45	28.99	30.08	36.20	46.43	47.23
광주	44.31	45.52	53.31	47.46	52.90			39.51	27.52	27.26	29.72	36.67	45.96	40.84
대구	52.64	51.09	53.45	48.26	50.47			40.07	32.63	29.44	30.67	37.08	46.80	48.74
대전	52.59	52.00	55.93	49.05	52.83			38.17	27.19	25.22	27.83	35.35	47.28	47.05
부산	47.28	48.69	53.20	52.27	57.42			43.68	38.05	35.55	32.86	36.83	44.88	43.50
서울	55.30	56.06	60.44	53.04	56.35			39.92	30.54	28.13	27.86	35.03	45.65	48.63
세종	59.50	54.70	60.60	50.60	51.50			37.18	27.45	27.83	28.50	34.25	51.25	51.13
울산	45.75	47.75	53.14	53.47	58.57			45.04	39.76	34.04	32.30	36.52	44.30	40.45
인천	55.45	57.53	61.97	55.77	58.83			42.91	34.29	33.59	34.08	40.55	49.33	50.49
전남	40.27	42.58	45.75	41.61	47.05			35.06	28.19	27.96	26.66	30.45	37.39	34.39
전북	56.51	57.49	62.36	54.82	58.58			45.57	29.28	29.05	32.35	41.18	52.06	48.75
제주	41.88	44.18	51.42	47.91	52.70			35.33	28.48	28.88	29.97	36.15	44.15	37.60
충남	51.80	52.31	58.45	50.35	52.99			38.41	27.92	28.49	29.88	38.07	46.99	43.00
충북	62.88	61.47	64.85	54.97	56.04			41.77	29.06	26.24	29.94	40.50	55.28	56.01

[시각화를 위한 HINT]

1. “기준년월”의 “월”을 열 선반에 놓아주세요.

- Window : 기준년월에서 마우스 오른쪽 버튼을 클릭하고 열 선반으로 드래그
- MAC : Option 키를 누르고 기준년월을 열 선반으로 드래그

위쪽에 있는 (불연속형) “월”을 선택해 주세요.



2. “시도”를 행 선반에 놓아주세요.

3. 마크의 모양은 사각형을 선택해 줍니다.

4. 마크의 색상을 평균 PM10 수치로 표현해보세요.

미세먼지가 높을 수록 붉은색으로 미세먼지가 낮을수록 파란색으로 표현해주세요

3. 테이블 계산식을 이용한 시간 분석

빅맥 지수는 각 나라의 물가, 구매력 평가를 비교하기 위해 고안된 경제지표입니다. 나라 별 빅맥 가격을 달러로 환산하고 이를 비교하면 나라 별 물가를 비교할 수 있습니다.

big-mac-rw-index 데이터에 있는 Dollar Price는 달러로 환산된 빅맥 가격입니다.

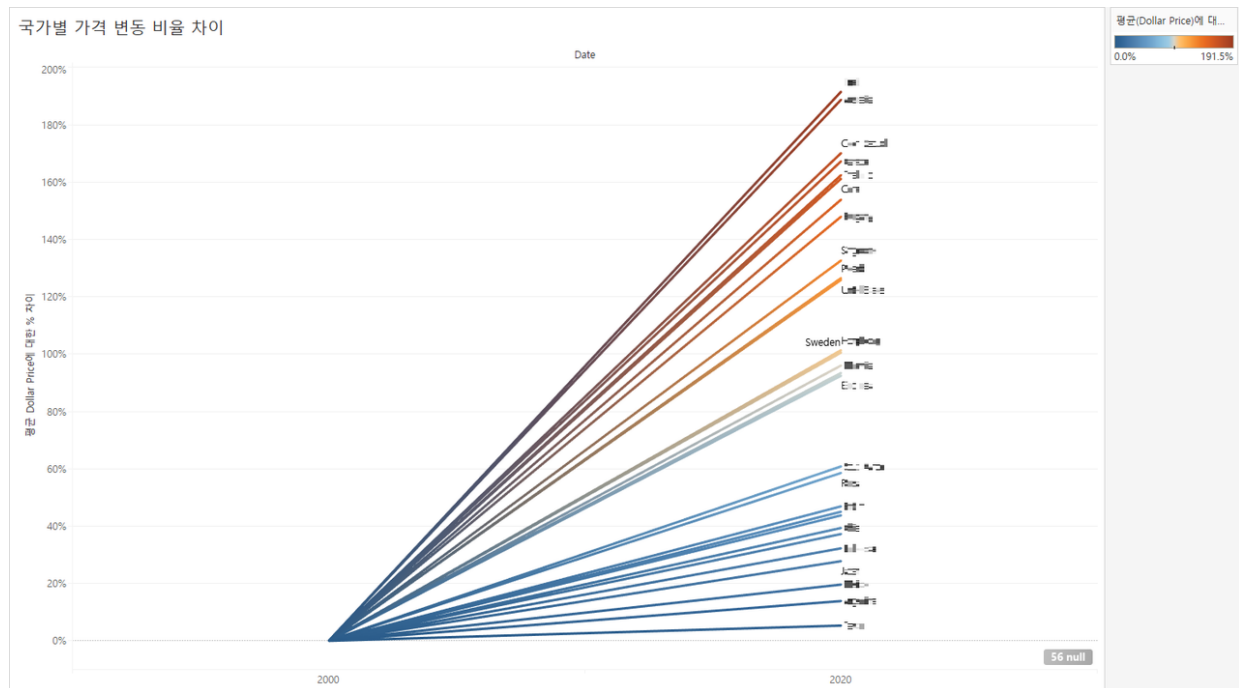
Dollar Price를 이용해서 2000년도 대비 2020년도의 빅맥 가격이 국가별로 얼마나 상승했는지 비교해보려고 하는데요.

그런데 여기서 잠깐!!

빅맥 가격 상승 폭을 비교할 때 단순히 가격으로 비교하는 것이 아니라 **가격 상승 비율**을 통해서 어느 나라 빅맥 가격이 가장 큰 폭으로 상승했는지 살펴봅시다!

예를 들어, A국가에서는 2000년도에는 500원 이던 빅맥이 1000원으로 올랐다면 100% 증가, B국가에서는 1600원 이었던 것이 2000원으로 올랐다면 25% 증가했겠죠. 그렇다면 A국가의 빅맥 가격이 작더라도 A국가는 더 많은 가격이 상승한 것을 볼 수 있을텐데요. 이와 같은 기준으로 아래와 같이 시각화를 나타내볼까요?

사용 Data = big-mac-rw-index



시각화 후 아래 질문에 답해 보세요!

Q: 어느 나라 빅맥 가격 변동 비율이 가장 큰가요?

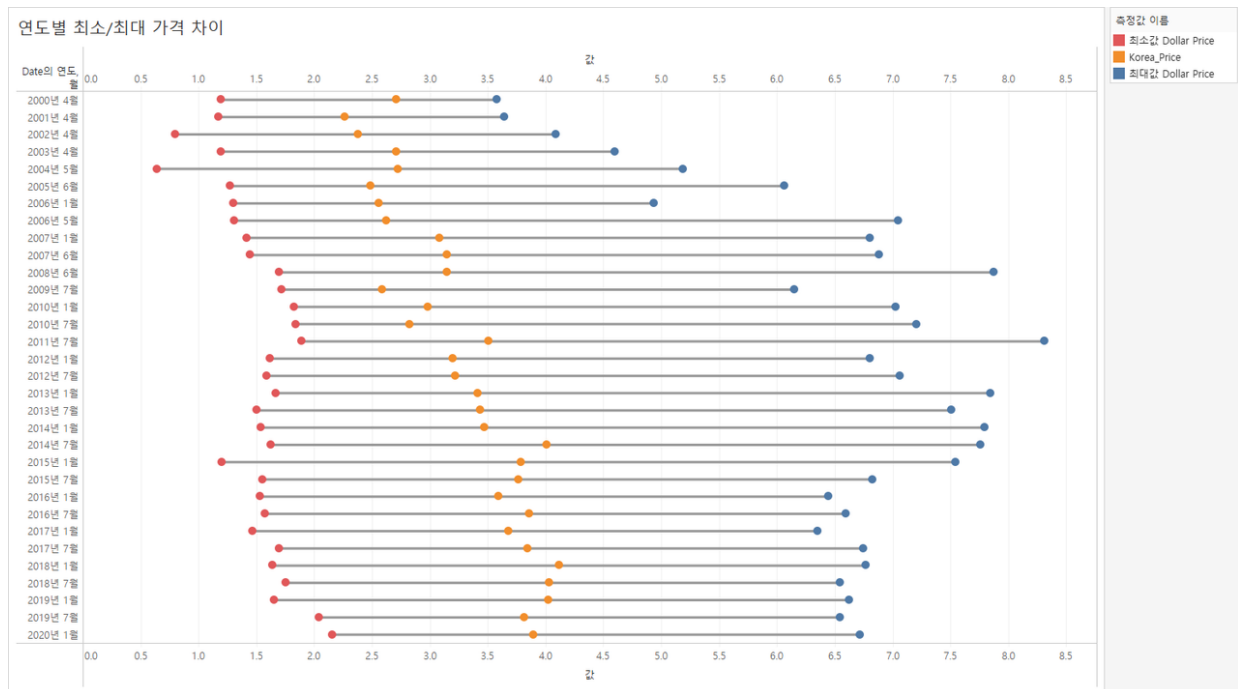
[시각화를 위한 HINT]

1. Date 필드를 옆에 가져다 놓은 후 **2000년도와 2020년 값만 유지**하세요.
Ctrl 키를 이용해서 2000년 마크와 2020년 마크를 선택하고 “**이 항목만 유지**”를 써보시는 것도 좋을 거 같네요.
2. Dollar Price에 대해 “**퀵 테이블 계산**” 중 “**비율 차이**”를 사용해 보세요.
값이 제대로 나오지 않는 것 같다면 3번을 고려해 주세요!
3. “**테이블 계산 편집**”을 이용해서 테이블 계산식의 **기준**을 변경해 보세요. (2000년은 데이터의 가장 “**첫 번째**” 값입니다.)

4. 덤벨차트

덤벨차트를 이용해서 **연도 별로 최소 빅맥 가격과 최대 빅맥 가격의 차이 변화**를 한 눈에 살펴봅니다. **Korea_Price**를 이용해서 한국의 빅맥 가격 변동도 함께 살펴보면 더 좋을 것 같네요. (*Korea_Price는 어떻게 계산했는지 편집 버튼을 이용해서서 살펴보세요)

사용 Data = big-mac-raw-index



[시각화를 위한 HINT]

1. “**행**”에 “**년월**” 기준으로 알약을 가져다 놓으세요.
마우스 오른쪽 버튼을 클릭한 채로 Date를 행에 가져다 놓으시고, (**불연속형**) “**년월**”을 선택하시면 되겠죠?
2. 여러 개의 측정값을 하나의 축에 표현하기 위해서는 “**측정값**”을 이용해야 합니다. (“측정값 이름 및 측정값” 동영상상을 통해 두 개념은 알고 계실 것이라 생각합니다!!)
“**열**”에 “**측정값**”을 가져다 놓고, “**측정값 이름**”을 필터로 걸어 “**Dollar Price**”와 “**Korea_Price**”만 체크해 주세요.
3. 측정값 선반에 있는 “**Dollar Price**”의 집계 형식을 “**최소값**”으로 변경해서 빅맥 최소 금액을 찾아 보세요.

(선반에 있는 Dollar Price에서 마우스 오른쪽 버튼을 클릭해서 변경하면 됩니다.)

4. 측정값 선반에 “Dollar Price”를 한 번 더 가져다 놓고 집계 형식을 “**최대값**”으로 변경해서 빅맥 최대 금액을 찾아 보세요.
(측정값 선반에 최소값(Dollar Price), 합계(Korea_Price), 최대(Dollar Price)가 제대로 들어갔나요?)

5. 마크의 **형태**를 “**원**”으로 표현하고, 마크의 **색상**은 “**측정값 이름**”으로 구분해 봅니다.

6. 옆에 있는 “측정값”을 옆으로 복사해서 2개의 측정값이 보이도록 만들어 보세요.
Ctrl키를 누르고 측정값을 클릭한 상태로 그 옆으로 드래그 해보세요~

7. 복사된 측정값의 마크 **형태**는 “**라인**”으로 변경하세요. 색상에 있는 “**측정값 이름**”을 “**경로**”로 옮겨보세요.

8. 2개의 축을 이중축으로 만들어 보세요. 이중축 후 축을 동기화 시켜 주세요.

9. 선이 위로 올라오나요? 옆에 있는 2개의 측정값 위치를 서로 바꿔보세요~