

공공데이터 사이언스 Spotfire 사용방법

● 다크모드 또는 화이트 모드

[Visualizations]-[Canvas styling]-[Dark]

● 언어 바꾸기

[tool]-[Option]-[Application]-Language

● 데이터 로드시 초기 시각화

[tool]-[Option]-[Document]-initial visualization ...

● 데이터 로드 방법

- 1) Drag & Drop
- 2) [Data]-[Add Data]-[Browse local file]
- 3) [File]-[Open]-[Browse local file]
- 4) (+)-[Browse local file]
- 5) 원하는 Data를 드래그 한 후 Ctrl+c Ctrl+v로 가져오기

● 저장하기

데이터 로드 설정 대화상자에서 해당 설정을 잘 선택해야 한다.

- 1) Always new data(항상 새 데이터)
- 2) New data when possible(가능한 경우 새 데이터)
- 3) Stored data(저장된 데이터) -> *(과제나 프로젝트를 저장, 제출할 때 이것으로 저장)*

● 데이터의 형식 바꾸기

잘 안됨. 확인 해봐야 함.

● 마킹

table 형식은 기본적으로 마킹 설정이 되어있다.

table 제외한 다른 형식은 수동으로 마킹 설정을 해줘야 한다.

(해당 차트 우 클릭)-[Properties]-[Appearance]-Use separate color for marked items 체크

● 마우스 드래그로 마킹

Alt + 마우스 좌클릭

Ctrl은 추가할 때 사용

● 전체 마킹 지우는 방법

- 1) [Edit]-[Marked rows]-[마크 취소 해제]
- 2) 빈 영역 클릭
- 3) Ctrl + e

● 내보내기

[file]-[Export]-[사진, pdf, ppt등 클릭]

● ppt로 내보낼 때, 각 상황을 케이스로 나눠서 출력할 때

(화면 우클릭)-[Propertile]-[Trellis]-[Panels] 클릭-[Manual layout] 클릭- 원하는 행, 열의 크기 설정

● 필터링

- 1) [View]-[Filters]
- 2) 우측 상단의 마크 클릭

● 필터링 초기화

- 1) Filters 우클릭 - [Reset All Filters]
- 2) Filters 우측에 있는 마크 클릭

● 필터의 이름 재정의

(Filters 우클릭)-[Rename]-이름재설정

● 필터에 표기하는 행 편집

(Filters 우클릭)-[Organize filters]-[show/Hide]-조작

● 필터링 스키마

(Filters 우클릭)-[Show Filter scheme Menu]-(Filter scheme 우측 아래 화살표 클릭)-[New Filtering scheme]
-이름생성 후 필요한 필터링을 설정한다-추가적인 필터링은 [New Filtering scheme]를 다시 생성하고 재정의 한다.

● 생성한 필터링 스키마를 차트에 적용

(해당 차트 우클릭)-[Properties]-[Data]-[Limit data using filterings]-해당 스키마 선택

=====

테이블 차트

● 테이블의 설정 열기

- 1) (테이블차트를 키고 우측 맨 위에 톱니 바퀴를 누르기)
- 2) (차트 아무대나 우클릭)-[Properties]

● 테이블 제목설정

(차트 아무대나 우클릭)-[Properties]-[General]

Title : 차트의 제목 변경

show title bar : 체크를 해제하면 차트의 제목을 사라지게 할 수 있다.

Description : 주석 작성

테이블의 제목을 잡고 드래그하면 원하는 위치에 차트를 배치할 수 있다.

● 테이블 데이터 설정

(차트 아무대나 우클릭)-[Properties]-[Data]

Marking : 현재 생성된 마킹 종류 중 하나를 선택하여 마킹을 적용할 수 있음

Limit data using marking - 해당 마킹 종류의 체크 : 해당 종류의 마킹이 적용된 데이터만 보여줌.

Limit data using marking - New : 새로운 형식의 마킹을 생성할 수 있음.

Limit data using filterings - 해당 스키마를 체크 - 필터 스키마를 클릭하여 적용할 수 있음

Limit data using filterings - New : 필터 스키마를 새롭게 생성할 수 있다.

Limit data using expression - Edit... - 조건입력 : 조건에 부합하는 데이터만 출력할 수 있게 함

● 테이블 모습 설정

(차트 아무대나 우클릭)-[Properties]-[Appearance]

Header row height : 제목의 높이 설정

Data rows height : 애트리뷰트(행)의 높이 설정

Number of frozen columns : 컬럼의 위치 고정

Use separate color for marked items - 체크해제 : 마킹이된 애들만 선명하게 보임

● 테이블 폰트 설정

(차트 아무대나 우클릭)-[Properties]-[Fonts]

● 테이블 칼럼 설정

(차트 아무대나 우클릭)-[Properties]-[Columns]

Available columns - Selected columns : 보고싶은 열만 선택하여(select) 표시한다.

● 테이블 색상 설정

(차트 아무대나 우클릭)-[Properties]-[Colors]

Add : 색상으로 표시하고 싶은 열을 추가한다.

(표시하고 싶은 열을 클릭) - (열의 데이터가 숫자형인 경우) - color mode : 각 설정에 맞는 표시를 해준다.

- Gradient : 숫자들의 값들로 숫자의 값이 크면 진하도록 그라데이션 표기

- Segments : 구간에 따라 같은 색으로 애트리뷰트 표시

- Fixed : 같은 색으로 값이 있는 애트리뷰트에 표시

- Unique value : 값들의 종류에 따라 다른 색상을 표기

- Apply color to - Background : 색을 배경에 적용

- Apply color to - Text : 색을 문자에 적용

Add Point - 설정 : 설정한 값에 대한 점을 표시하여 그룹 표기

Add Rule... - 설정 : 설정한 규칙에 대하여 해당 열에 색상을 적용하는 방법

● 테이블 순서 설정

(차트 아무대나 우클릭)-[Properties]-[Sorting]

1) Sort by : 하나 또는 중첩의 정렬을 실행할 수 있음

2) (칼럼 이름 좌클릭) - Sorting : 하나의 정렬을 저장할 수 있음

● 테이블 범례 설정

(차트 아무대나 우클릭)-[Properties]-[Legend]

Show legend : 범례를 표시의 유무를 표기

Display the folling legend items : 범례에 포함시킬 내용을 선택

● 테이블 필드 및 애트리뷰트 설정

(해당 필드 좌클릭)

순서변경, 너비, 정렬, 배경 및 텍스트 채우기, 열 숨기기 등을 설정가능

열을 숨겼을 시 [설정]-[열]에서 다시 추가할 수 있음

=====

막대 차트

● X축과 Y축

X축은 웬만하면 이산형 자료가 합당하다.

Y축은 집계를 꼭 설정해야 한다.

(+)버튼을 통해 추가적인 그룹을 만들 수 있다.

● 막대차트의 모습

(차트 아무대나 우클릭)-[Properties]-[Appearance]

Orientation : 가로, 세로 막대 설정

Layout : 각 설정에 맞는 레이아웃을 설정

- Side-by-side bars : 옆에 표기

- Stacked bars : 쌓아서 표기

- 100% stacked bars : 비율을 맞춰서 표기

Bar width : 막대 너비

Transparency : 막대 투명도

Show shadows indicating the unfiltered data : 원본데이터는 그림자로 표기하고 필터링 된 데이터 상태는 채색으로 표기해줌

Use separate color for marked items : 막대 차트에서 마킹의 유무 선택

Sorting : 순서배열

- Sort bars by value : 내림차순으로 막대 정렬

cf) [Category Axis]-[Scale]-[Reverse scale] 과 같이 설정하면 올림차순으로 막대 정렬

- Sort bar segments by value :

- Reverse bar segment order :

● 막대차트의 형(단위)변환

(차트 아무대나 우클릭)-[Properties]-[Formatting]

X축과 Y축 중 하나를 선택하여 각 축의 포맷을 정해준다.(글꼴, 형식 등)

● 막대차트의 X축 설정

(차트 아무대나 우클릭)-[Properties]-[Category Axis]

Columns : X축의 구분되는 열을 추가할 수 있다.

Scale labels - Show labels : 막대마다의 레이블 표기를 설정

- Horizontally : 수평으로 레이블 작성

- Vertically : 수직으로 레이블 작성

- Max number of labels : 레이블의 최대 문자수를 설정

Scale - Reverse scale : 막대의 순서를 반대로 뒤집음

● 막대차트의 Y축 설정

(차트 아무대나 우클릭)-[Properties]-[Value Axis]

Columns : Y축의 구분되는 열을 추가할 수 있다.

Single scale : Y축에 해당되는 데이터의 종류가 1개일 때

Dual scale : Y축에 해당되는 데이터의 종류가 2개일 때

- 2번째 카테고리를 오른쪽에 배치하려면 해당 카테고리를 스케일을 오른쪽으로 옮긴다.

Multiple scale : Y축에 해당되는 데이터의 종류가 3개 이상일 때

Scale labels - Show labels : 막대마다의 레이블 표기를 설정

- Horizontally : 수평으로 레이블 작성

- Vertically : 수직으로 레이블 작성

- Max number of labels : 레이블의 최대 문자수를 설정

Scale

- To the left : Y축을 왼쪽에 생성

- To the right : Y축을 오른쪽에 생성

- Reverse scale : Y축 범위를 뒤집어서 막대의 방향을 위에서 내려오도록 설정

Range

- min과 max에 값을 대입하여 사용

● 막대차트의 색상 설정(누적 막대 그래프 표시할 때도 포함)

(차트 아무대나 우클릭)-[Properties]-[Colors]

(표시하고 싶은 열을 클릭) - (열의 데이터가 숫자형인 경우) - color mode : 각 설정에 맞는 표시를 해준다.

- Gradient : 숫자들의 값들로 숫자의 값이 크면 진하도록 그라데이션 표기

- Segments : 구간에 따라 같은 색으로 애트리뷰트 표시

- Fixed : 같은 색으로 값이 있는 애트리뷰트에 표시

- Unique value : 값들의 종류에 따라 다른 색상을 표기

Add Rule... - 설정 : 설정한 규칙에 대하여 해당 열에 색상을 적용하는 방법

Columns-누적으로 쌓고 싶은 기준 설정

상위 몇 개 혹은 하위 몇 개의 막대만 색상을 다르게 주는 방법

Columns-Value axis values-Add Point 혹은 Add Rule로 설정-각 설정의 순서를 배치하여 누적으로 색상을 여러개 표시가능

● 막대차트의 라벨 설정

(차트 아무대나 우클릭)-[Properties]-[Labels]

Show label for : 각 해당하는 옵션으로 라벨을 부여

- All : 모든 막대에 라벨링
- Marked rows : 마킹을 한 막대에 라벨링
- None : 하지 않음.

Type of labels : 각 해당하는 옵션으로 라벨을 부여

- Complete bar : 막대 위쪽에 표기
- Bar segments - Value : 중앙에 값으로 표기
- Bar segments - Percentage : 중앙에 퍼센트로 표기

Label orientation

- Horizontally : 수평으로 작성
- Vertically : 수직으로 작성

● 막대차트의 툴팁 설정

(차트 아무대나 우클릭)-[Properties]-[Tooltip]

Display the following values : 표시한 데이터를 막대에 커서를 가져다 대면 쪽지 형식으로 볼 수 있게 설정함

● 막대차트의 Trellis 설정

(차트 아무대나 우클릭)-[Properties]-[Trellis]

Rows and columns : 각 기준에 맞는 형태로 격자 표시를 한다.

- Rows : 행 자르기
- Columns : 열 자르기
- Pages : Page로 자르기

Panels : 자동으로 잘라서 표기를 해준다.

-Manual layout : 수동으로 가로와 세로의 수를 입력하여 격자로 잘라서 표기

● 막대차트의 선과 곡선 설정

(차트 아무대나 우클릭)-[Properties]-[Lines & Curves]

Add-Horizonta Line-straight Line-조건 : 조건에 맞는 수평 직선 설정을 추가한다.

Add-Vertical Line-straight Line-조건 : 조건에 맞는 수직 직선 설정을 추가한다.

● 막대차트의 숨김과 보기 설정

(차트 아무대나 우클릭)-[Properties]-[Show/Hide items]

Add ...-Column-Rule type : 선택한 필드의 조건의 해당하는 데이터만 한정적으로 출력

● 해당 막대에 대한 상세 데이터 막대그래프 표시

우클릭-[Create Details visualization]-bar chart-

● 막대 그래프에서 막대 크기에 따른 정렬

우클릭-[Sort bar by Value]

● 막대들 중 원하는 막대만 골라서 표현하는 방법

[속성]-[Show/Hide items]-Add-원하는 계산을 설정, 누적 설정 가능하다.

● 막대 값이 0인 것을 제외시키는 방법

필터에서 그 값에 대한 컬럼을 찾는다.-그 컬럼의 범위를 최소 0.01로 변환한다.(더블클릭해야 바뀜)

● 막대 그래프로 연속형 데이터를 넣으면 막대가 얇고 길게 된다. 이것을 이산형처럼 바꾸는 방법
x축 아래 우클릭-[Categorical Scale]

● 보고 싶은 데이터의 새로운 차트 열기
(보고싶은 데이터 우클릭)-Create Details Visualization-(원하는 차트 선택) : 보고싶은 데이터만 사용한 차트 생성

● 뷰
[View]-[Bookmarks]
현재 필터나 클릭을 통해 만들어진 데이터 차트 모양을 그대로 저장하기 위해서 사용
(저장하고 싶은 부분에서)-[View]-[Bookmarks]-이름작성-(+)

● 우리가 지정하는 필드, 태그 - 다만 저장이 안되므로 키 값을 이용해 새롭게 표를
[View]-[Tags]-(만들기 클릭)-tag이름작성-주석-New-tag에 들어갈 필드를 작성
[Properties]-[Lines & Curves]를 통해 차트에 선과 곡선을 그리고
해당하는 부분의 데이터를 선택한 다음 tag에 만들어놓은 필드에 저장
[Data]-[Data Table Properties]-[Key columns for linked data]-새로운 키 값을 만듦
[Data]-[Transform data]-[Insert]-키를 불러오거나 새로 만들어서(RowId())를 사용하면 만들어짐 차트에 저장
키 값이 될 수 있는 값들을 만들거나 선정하여 tag필드를 차트에 포함시켜 저장 해놓는다.

산점도 그래프

- 데이터를 좌표평면 상에 마커로 찍어서 표현
- 두축, 마커가 의미하는 바가 무엇인지 명확하게 정의되어야 함
- 회귀분석을 할 때 매우 중요한 차트
- 막대 그래프는 y를 표현할 때 집계방법을 필수적으로 물어보지만
- 산점도 그래프의 y를 표현할 때는 집계방법을 따로 물어보지 않는다.

● 산점도에서 어떤 목적의 가장 많이 분포되어 있는가를 확인할 때 투명도를 사용한다.
(차트 아무대나 우클릭)-[Properties]-[Appearance]-[Transparency]
- color by를 삭제하고 보는 것을 추천

● 변수가 이산형이라 값이 한줄로 표기가 될 때 값들의 차이를 알아보기 위해 지터링을 사용한다.
(차트 아무대나 우클릭)-[Properties]-[Appearance]-[Jittering]
Use the same amount if jittering for both X and Y : 체크를 풀면 지터링 할 때, x, y를 따로 움직이게 한다.
이산형 변수를 값 사이에 공간이 많아 표현하기 적당한데
연속형 변수를 값 사이에 공간이 적어 표현할 때 이웃 값의 범위를 침범할 수 있으므로 주의 필요
이산형 * 이산형 데이터를 할 때 x, y 지터링을 잘 활용 해보자.

● size by, color by, shape by 를 이용하여 데이터를 추가적으로 표현할 수 있지만,
되도록이면 복잡하지 않게 3개 정도까지만 사용하자.
- 여러 가지를 표현하고 싶다면 trellis를 사용해서 격자로 표현하자

● 이산형 * 이산형 데이터에서 각 카테고리별 데이터의 수를 비교하는 방법

(모든 설정 지우고)-(color by에 no 함수를 사용)-색이 짙은 것이 많은 데이터가 있음을 뜻한다.

● 산점도 모양을 타일 모양으로 바꾸는 방법

(차트 아무대나 우클릭)-[Properties]-[Shape]-[Tiled markers] : 파일 모양으로 데이터의 양을 비교. 레이블을 같이 쓰면 좋음
이산형 * 이산형 자료의 수량 비교를 파악하는데 잘 사용될 것 같다.

● 데이터의 라벨 붙이기

(차트 아무대나 우클릭)-[Properties]-[Labels]-Label by 에 기준으로 할 변수를 입력

-show labels for: 라벨을 표기할 데이터 선택

all : 모두 라벨 표시

Marked rows : 마크한 데이터만 라벨 표시

None : 없다.

● 버블차트 만들기

- 이산형 * 이산형 데이터로 클을 잡고 연속형이나 이산형의 자료를 by를 통해 표시하여 사용

- 점을 파이차트로 하여 추가적으로 사용할 수 있음

(데이터를 배치하고 모든 설정을 초기화한다.)-(마크 by 설정을 하고)-[Properties]-[Shape]-[pies]-[Sector size by] 설정

**** 중요, 파이차트는 크기가 큰 것부터 정렬하여 시계방향으로 표시해야한다.

[Properties]-[Shape]-[pies]-[Sort sectors by size]를 클릭하면 정렬이 이루어진다.

- 버블의 크기를 키우고 싶으면 [Properties]-[Size]에서 설정

*** 여기서도 범위마다 차이를 확인하고싶다면 trillis를 잘 활용하자.

● 산점도에서 데이터의 값들이 표시되는 순서를 선택하려고 한다.(디폴트는 입력한 순서대로 그려짐)

[Properties]-[Drawing Order]-변수설정

● 라벨링

[Properties]-[Labels]-변수설정

● Mark by, color by 등등은 고른 데이터의 종류당 하나의 데이터만 표시하겠다는 정보를 담고 있다.

● 상관관계를 표시하는 방법

알스퀘어 값이 0에 가까우면 '변수 사이에 상관성이 없다.'이고, 1에 가까우면 "두 변수 사이에 상관성이 매우 높다."라고 한다.

알스퀘어 값이란 한쪽 변수가 한쪽 변수를 설명할 비율로써, 키와 몸무게의 R^2 이 0.389이라고 한다면

'1000명 중에 389명 정도 키를 통해서 몸무게를 설명할 수 있다.'라고 하거나

'키가 몸무게를 또는 몸무게가 키를 1차원적으로 설명할 수 있는 비율로 39%정도 설명할 수 있다.'라고 해석이 가능하다.

R 값이 0.5이상이면 완벽할 텐데 쉽지 않다. 실제로 0.5 이상은 거의 없으며, 0.4만 넘어도 높다고 평가한다.

[Properties]-[Lines & Curves]-straight Line Fit : Straight line에 체크-[Label and Tiiltip]- R^2 Label에 체크

● 3D 플랏

- 표현할 때 쓰긴하는데 효율성이 떨어진다.

- 차라리 2차원과 trillis를 사용하는 것이 더 시각적으로 보기 좋다.

● 파이차트

- 가장 추천하지 않는 차트

- 12시를 기준으로 영역이 가장 큰 값부터 차례로 표기하는 것이 정석이다. 정렬을 해야 감점이 없음

- Sector size by를 표현하지 않는 경우가 있는데, 꼭 해줘야 감점을 당할 일이 없다.

- 정렬하는 법

[Properties]-[Appearance]-Sort sectors by size 체크

- Sector size by 하는 법

[Properties]-[Size]=Sector size by에서 설정

여기서 Sector과 Size를 설정할 수 있는데 기본적으로 같은 값을 가져가는 경우가 많지만.

전문적으로 다루거나 나타내야 할 값이 많으면 다르게 설정하여 데이터를 구분할 요소로 가져갈 수 있다.

여기서도 Trillis를 잘 활용하면 너무 좋다.

● 라인차트

- 시간에 따른 값의 변화를 보기위해 많이 사용

- x값에 날짜 데이터가 들어가면 화살표 표시가 생기면서 사용할 날짜의 종류를 선택할 수 있게 된다.(M,Y, S 등등)

- 만약 이 화살표가 생기지 않으면 날짜 데이터로 인식을 못하는 경우이다.

- line by, color by등으로 나타내는 선들은 기준에 따른 분류값의 종류대로 선들이 생성된다.

[Properties]-[Appearance]-[Marker size]를 통해 라인 상 기준값들의 마커 크기를 조절할 수 있다.

- y축의 범위를 조절하는 방법

[속성]-[Y-axis]-min과 max에 값을 대입하여 사용

● 트리맵

- 금융가에서 내가 가진 포트폴리오가 뭐가 문제인지, 지금 하고 있는 투자가 어떻게 되가고 있는지 큰것부터 차례대로 확인하는데 유용하다.

- 가장 큰 사각형이 왼쪽 상단 모서리에, 가장 작은 사각형이 오른쪽 하단 모서리에 배치되어야 한다.

- Hierarchy(계층)가 중요하다.

- 계층 구조로 나타내는 어떤 데이터를 표현할 때 탁월하다.

- 색상은 예쁘게 하는 용도로 자주 사용된다.

- Hierarchy(계층)을 먼저 주고 size와 color순서대로 설정한다.

- 보편적으로 사이즈는 양적 데이터를 넣고, 색상은 질적 데이터를 넣는다.

- 권역을 여러 개 줄 수가 있다.

- 권역으로 나눈 카테고리 중 더 자세히 보고 싶은 곳이 있으면 이름을 더블클릭하면 된다.

- 뒤로 가고 싶으면 All을 클릭하면 된다.

- 데이터들의 자세한 값을 보고 싶다면 마우스 커서를 위에 올려놓으면 라벨이 보여진다.

- 계층 슬라이드 만들기

Hierarchy를 마우스 우클릭-Create Hierarchy

- 계층 슬라이드 만들었을시 계층 수정 및 이름 수정(계층 슬라이드를 만들면 초기 이름이 따로 지정하지 않는 이상 계속 유지)

[Data]-[Column Properties]-[Column and hierarchy]-맨 밑에 추가한 권역 클릭-Edit.. 클릭-계층 수정-이름수정

- 컬러를 설정할 때 각 분류를 다른 색상으로 하고 싶다면 UniqueConcatenate로 설정하라

color by-Aggregation-UniqueConcatenate

- 어떤 값에 측정 색상을 넣어주려면

[속성]-[color]-[Gradient]-[Add Point]로 작성

- 여러 속성을 같이 보고 싶으면 treills를 잘 사용하자
-

● 크로스 테이블

- 일반적으로 엑셀의 피벗테이블과 같다고 보면된다.

- 원하는 만큼의 데이터를 ctrl+c 한 다음 ctrl+v로 붙여넣으면 보고 싶은 데이터만 볼수 있다.

- 데이터의 단위라든지 Format을 바꾸려면

[속성]-[Format]-[Category]-Currency 화폐 등등 설정하여 사용

- y축에 여러 기준을 두고 싶다면

y축 선택 옆에 +기호 누르고 원하는 기준 추가하기

- 색상 속성에서 원하는 기준에 색상을 설정하거나 룰을 추가하여 상위 혹은 하위 데이터에 색상을 지정할 수 있다.

- 계층이 여러개 있는 곳에는 계층 슬라이드를 만들 수 있다.

- 데이터 테이블로 내보내기 위해서는

[File]-[Export]-[data to file]-[data table]

- 데이터에 대한 행 혹은 열의 총 합계를 표시하는 방법

우클릭-[Grand Total for Row] 혹은 [Grand Total for Columns]

- 정렬은 열의 이름을 클릭하여 간단한 아이콘으로 설정한다.

● 콤비네이션 차트

- 막대 그래프와 선 그래프의 기능을 결합하는 시각화이다.

- 막대 그래프는 color by가 디폴트이고 콤비네이션 차트는 series by가 디폴트이다.

- 막대와 선그래프를 함께 표시하기 위해서 y축 범례에 계층을 2개로 설정한다.

- 왼쪽 필터의 Series by에서 해당 범주의 아이콘을 클릭한 후 바, 라인 차트 중 하나를 고른다.

- 범위나 단위에 차이로 갭이 커서 높이 차이가 많이 난다면

y축 우클릭-Multiple Scale

- 라인 차트와 바 차트가 만나는 곳에 마크를 하는 방법

[속성]-[Appearance]-[show line markers]-크기 조절

- 막대 그래프의 너비 조절

[속성]-[Appearance]-[Bar width]-크기 조절

- 하나의 속성을 표현하고 싶다면 y축 계층에 하나더 추가한다.

- 포맷 설정을 하여 값의 표시를 하려면

[속성]-[Formattion]-원하는 데이터를 클릭-Category에서 설정

- 라벨링을 달기 위해서는

[속성]-[Labels]-[show labels for]-all 혹은 마크 범위로 설정

- y축의 범위를 설정하는 방법, 물결 표시 느낌처럼

[속성]-[Y-axis]-[Scale Range...]-min과 max에 값을 넣어주면 된다.

● 폭포 차트(워터폴 차트)

- 보통 제조업체에서 사용을 하며, 금융쪽에서 손해부분을 정확하게 보기위해서 사용한다.

- 보통 빨간색이면 감소, 파랑색, 초록색은 증가라고 표시된다.

- 보통 x축에 타임시리즈를 넣는다.

- 타임시리즈에 일-달-분기-년처럼 분기, 년이 포함된다면 분기별 년도별 종합 표시를 할 수 있다.

x축에 분기가 포함되는 형식으로 지정-[속성]-[Total Bars]-show total bars-one total bar after each value in-quarter 클릭

- y축에 단위 변경 및 표기나 각 막대의 값을 표기하기 위해서는 라벨링이랑 포맷팅을 잘 사용한다.

● 요약 테이블(써머리 테이블)

- 통계 정보를 테이블 형식으로 요약한 시각화이다.

- 행을 잡는 기준은 숫자만 자동으로 잡는다.

- 문자도 수동으로 추가 시킬수 있는데 카운트만 표기할 수 있다.

- 아래의 두 수치를 동시에 사용하면 중복값이 있는지 확인가능하다.

UniqueCount : 중복값을 제외한 값의 count

count : 중복값을 포함한 값의 count

[속성]-[Statistical Measures]-UniqueCount와 count를 추가

- 이산값있는지 확인

Avg와 Mid를 비교하면 알 수 있다.

- y축에 계층을 추가하려고 한다면

[속성]-[Columns]-Categorization-+ 기호를 누르고 계속 추가한다.

● 그래픽 테이블

- 있어보이는 차트로 꾸밀 때 사용하는 차트이다.

- 있어보이게 꾸미는 테이블이다.

- 동적항목(스파크라인, 계산된 값, 조건부 아이콘 및 불릿 그래프)과 함께 칼럼 표시

- line차트와 비슷하게 보이는 Sparkline이 존재한다.

- 계층을 추가하여 여러개를 표시할 수 있다.

y축에 + 표기를 눌러 속성을 추가한다.

- 스파크라인의 추가 및 속성을 설정 방법

[속성]-[Axes]-add-Sparkline

- 이름 설정

[General]-name 설정

- 스파크라인의 시작점과 끝점의 데이터 표시

[Appearance]-Show Y-axis values for-Starting Point 와 End point 체크 표시-Y-axis values width로 너비 설정

- 스파크라인의 라인 두께 설정

[Appearance]-Line width에서 설정

- 계산된 값의 추가 및 속성을 설정하는 방법

[속성]-[Axes]-add-Calculate value

- 계산된 값들의 부호나 형식 설정

[Formating]-Categorys에서 설정

- 특정값에 표시(상위 3개, 평균이상 등등)

[Valuies]-Add Rule-원하는 설정 등록

- 아이콘 추가 및 속성 설정하는 방법

[속성]-[Axes]-add-Icon

- 특정 값 혹은 범위에 값들에게 아이콘 표기하기

[Icon]-add Rule...-value,값설정-Color,색 설정-Shape.모양설정

- 불릿 그래프 추가 및 속성 설정하는 방법

[속성]-[Axes]-add-Bullet Graph

- 불릿 그래프의 특징

- 가로 막대는 데이터의 실제 값을 나타내고, 세로 막대는 데이터의 목표값을 나타낸다.

- 배경을 이루는 3가지 단계의 색상은 각각 불량, 보통, 양호에 범위를 알려주는 것으로 데이터가 어디에 속하는지 알려준다.

- 가로 막대 표시

[Bullet Graphs]-Calculate Values using-원하는 값을 설정, 색상도 color에 설정

- 세로 막대 표시

[Bullet Graphs]-Calculate comparative Values using-원하는 값을 설정, 색상도 color에 설정

- 배경 범위 설정

[Color Ranges]-show color range체크박스를 클릭하고-밑에 범위와 색상을 입력-추가적인 범위를 설정하고 싶으면 add로 추가

● 평행 좌표 그래프

- 심층 분석을 할 때 사용하는 그래프이다.
- 평행 좌표 그래프에서는 데이터 테이블의 각 행을 선 또는 프로파일로 매칭한다.
- 프로파일을 비교하여 유사성을 찾는 데 있어 평행좌표 그래프가 매우 유용하다.
- 컬럼마다 배열이 다르므로 컬럼간에 곡선의 높이를 비교하지 않는다.
- 각 컬럼의 데이터를 정규화 하여 최저 값을 0%로, 최대값을 100%로 간단히 표시하여 비교를 수행한다.
- 시계열을 나타내는 용도가 아닌, 선들의 패턴을 분석하는 차트
- LINE 그래프와 다르게 x에 여러 가지 변수를 표시하여 비교한다.

- 하나의 학교에 대해서 한 라인씩 지정하려는 생각으로 가장 먼저 Line by와 Color by를 조정한다.

차트 오른쪽에 ~~~by를 조정하는 곳에서 설정

-

[속성]-[Columns]-오른쪽에서 왼쪽으로 옮기면 x축에 변수가 추가되는 것이며, 여러 컬럼을 드래그하고 함수 계산을 하게 되면 동일하게 적용이 된다.

- 이상치 제거하는 방법 (전체 데이터의 y축 범위를 조정하기 위해서)

마우스 우클릭-[Marked rows]-[Filter out]

- 평행좌표그래프에서 각 요인 중 드래그한 데이터가 얼마나 많은 비중을 차지하는지 비교하기 위해 파이 차트와 함께 확인

마우스 우클릭-[Create Details Visualization]-Pie chart ... - Sector size by 꼭 함수와 함께 를 먼저 설정하고, color를 설정

- 같은 차트를 복제하기 위해서는

마우스 우클릭-[Duplicate visualization]

- 파이차트에서는 꼭 비중이 큰 값부터 차례대로 시계방향으로 누적으로 더해져야 한다.

[속성]-[appearance]-[Sort sectors by size]

- 평행좌표그래프에서 각 요인 해당 데이터를 정렬해서 보기 위해, closs table를 옆에 띄워서 확인한다.

마우스 우클릭-[Create Details Visualization]-closs table ... -정렬하고 싶은 기준 컬럼이름 클릭-해당 정렬수행 후 값 확인

● KPI 차트

- 있어보이는 차트로 꾸밀 때 사용하는 차트이다.
- 핵심 성과 지표를 나타내는 차트로 여러 KPI들을 타일 형태로 표현
- 부서별 그룹별 지표를 한번에 비교하고자 할 때 사용

- KPI 패널 너비 조정 및 마킹시 포인트 남기는 설정

[속성]-[Appearance]-Tile width not less than를 통해 너비를 조절할 수 있다.

[속성]-[Appearance]-Use separate color for marked items를 통해 마킹시 표시 설정을 할 수 있다.

- KPI 설정, Tile by, 비교 되는 값(따로 설정 못함), Value, x-axis까지 추가됐을 때 값 순서대로 패널에 표기가 된다.

[속성]-[KPIs]-add-[Values]- Tile by, Value, x-axis 차례로 값 설정

- 일정한 규칙에 의해 색으로 포인트를 주는 방법

[속성]-[KPIs]-add-[Color]-[Add Rule...]-평균, 상위, 하위 등등 설정

- 특정 값에 색으로 포인트를 주는 방법

[속성]-[KPIs]-add-[Color]-[Add Point...]-값이랑 색 설정

- 순서 정렬을 하고 싶으면 먼저 속성에 정렬을 지정한 다음 설정해야 best sort가 활성화가 된다.

[속성]-[KPIs]-add-[Sortion]-정렬 기준 설정-바탕에서 다시 마우스 우클릭-Sort Order-여기서 올림 내림 정렬을 설정

● 히트맵

- 심층분석, 계층적 자료에 사용하는 차트이다.
- 이상형 변수를 많이 비교할 때 사용할 수 있다.
- 색깔로 읽는 차트, 숫자 표기 불가하다.
- 숫자가 있으면 Scatter plot의 변형으로 사용한 것이다.

- 원하는 룰이나 점을 표시하여 색상을 표시한다.
- [속성]-[colors]

● 데이터 핸들링 - 분석 데이터

- 선택 데이터에 대한 자동 차트 추천 기능
 - 해당 컬럼에 대한 값 범위(연속형), 값 선택(이산형) 등을 조절 가능
- 해당 컬럼을 찾고-그 찾은 컬럼 이름 오른쪽 끝에 깔때기 모양 아이콘 클릭-여기서 설정한다.
- 컬럼의 이름을 재정의 할 수 있다.
 - 비어있는 값을 대체하는 것도 여기서 다뤄진다.
- 해당 컬럼을 찾고 선택한다.-톱니바퀴를 누르고 아래쪽으로 간다.-Empty values에서 빈 값을 특정 값으로 변경 가능
- 해당 열의 가지고 있는 데이터를 일정 기준에 의해 분리를 할 수 있다.
- 해당 컬럼을 찾고 선택한다.-톱니바퀴를 누르고 아래쪽으로 간다.-Split Column-기준, 최종열의 개수 등 설정

● 데이터 핸들링 - 컬럼 매치

- 비슷한 컬럼을 가지는 데이터를 추가하면 Spotfire가 알아서 두 테이블을 하나의 차트에 소화 해주는 기능이다.
- 비슷한 데이터를 불러온다-기준이 되는 데이터로 차트를 그린다-y축 + 클릭-데이터 클릭-비슷한 데이터 클릭-컬럼 설정
- 추가적으로 누적 막대, 나란히 막대, 100% 막대 설정 등은 차트 우클릭에서 해결한다.
 - 비슷한 테이블의 속성을 확인하는 방법
- [Data]-[Data Table Properties]

● 데이터 핸들링 - 데이터 다시 로드

- 원본데이터인 엑셀에서 값을 수정했을 때, 참조하는 데이터 차트에도 반영을 하기 위한 방법이다.
- 방법1. [Data]-[Reload all data]
- 방법2. [Data Canvas]-해당 데이터의 아래 팬케이크 같은 모양 클릭-맨 아래의 Data loaded at을 클릭

● 데이터 핸들링 - 데이터 캔버스

- 어떤 작업을 했고, 어떻게 변형 과정을 거쳤는지 도식화
- 데이터 변환 및 작업 내용 수정 가능
- 데이터 이름을 재정의 할 수 있다.

● 데이터 핸들링 - 데이터 교체

- 데이터 테이블의 양식(컬럼, 이름, 형식 등등)이 유사한 경우 데이터 교체 가능
- 데이터 테이블을 교체함으로써 만들어 둔 분석 화면을 그래도 이용 가능

- 단, 데이터를 교체해도 이름은 그대로이니 주의해야 한다.

데이터 캔버스-현재 데이터 클릭-왼쪽 아래 세로로 ... 클릭-Replace data source 클릭-바꿀데이터 선택 후 불러오기

● 데이터 핸들링 - 데이터 연결

- 두 개 이상의 데이터 테이블을 연결해주는 기능이다. Relation 적용 시 마킹과 필터링 연동

- 서로 다른 테이블을 연결시켜줄 기준 컬럼 필요

[Data]-[Data Table Properties]-Relations-Manage Relation-new-스포티파이어에서 알아서 기준 컬럼을 설정해준다.

- 테이블을 이어주는 컬럼은 기본키로 하면 좋는데 만약 복합키로 사용해야 한다면 여러개를 추가해서 연결한다.

[Data]-[Data Table Properties]-Relations-Manage Relation-new-원하는 컬럼을 기준으로 연결한후 추가

● 데이터 핸들링 - 데이터 값 바꾸기

- 해당 셀의 값 바꾸기

바꿀 셀 우클릭-Replace Value-All occurrences in column-바꿀값 입력-Apply

- 하나의 셀만 수정을 하고, 다음에 불러와도 프로그램이 바꾼 값으로 인식을 할 수 있도록 하기 위해서 키 컬럼을 설정해야 한다.

바꿀 셀 우클릭-Replace Value-this occurrence only 밑에 key columns 클릭-기준 컬럼을 선택했는데 안된다면 유일한 키가 이나라는 뜻이므로 유일한 키 컬럼을 만들어서 다시 설정해야한다.

다시 데이터 캔버스-왼쪽 아래 세로 ... 클릭-Transform data-Calculate new column-insert-rowid로 컬럼추가-이 행으로 위에 키 컬럼 설정-하나의 셀 수정이 가능하다.

● 데이터 핸들링 - 계산된 컬럼 추가

- 기존 컬럼들 값들 간의 산술 및 논리 표현식을 사용하여 원하는 새로운 컬럼을 만들어서 사용할 수 있다.

● 데이터 핸들링 - 데이터 변환

- 분석하고자 하는 데이터가 때때로 가장 적합한 포맷이 아니거나 오류를 포함하고 있는 경우 최적의 결과를 얻기 위해

- 데이터 변환이 유용할 수 있다. 데이터가 로드 될 때, 된 후 둘다 적용 가능하다.

- 형변환, 소수점 변환, 단위변환 등등을 여기서 할 수 있다.

데이터 캔버스-중앙에 데이터가 이어져있는 중간 + 클릭-Transform data-Change data type-insert-컬럼 클릭-새로운 형식지정

● 데이터 핸들링 - 함 컬럼 추가

- 연속하는 값 또는 범주 별 값을 그룹화하는 기능,

- 이상치 완화, 결과 해석에 용이하다.

[Data]-[Add binned column...]-Column 고르고-Specific : 사용자지정, Even intervals : 균등간격을 선택하여 구간 나눔

- 다른 옵션들은 PPT에 있음

- 분류하지 전 데이터를 bar 차트로 그린 다음, color by로 만든 분류 기준을 설정하면 구간별로 막대가 색칠된다.

● 데이터 핸들링 - 행 추가

- 기존 데이터 테이블에 새로운 데이터 테이블을 행으로 삽입하는 기능

데이터 캔버스- 중앙+ -Add rows-추가할 행이 있는 데이터 클릭-ok하기전 setting for added rows 선택-Match columns-자

동 매칭된 컬럼에서 필요없는 컬럼 삭제 및 추가-Identify origin of rows-기준 행에 표시할 값을 입력하는 곳-다시 차트를 만들면 행이 추가되어 있다.

● 데이터 핸들링 - 컬럼 추가

- 기존 데이터 테이블에 새 데이터 테이블에 대한 컬럼 추가
- 두 데이터 테이블을 이어줄 key column 필요

데이터 캔버스- 중앙+ -Add columns-추가할 열이 있는 데이터 클릭-ok하기전 setting for added rows 선택-Match columns-키 컬럼에 대한 정보를 보여줌-Join settings-조인유형을 여기서 설정-다시 차트를 만들면 행이 추가되어 있다.

● 데이터 핸들링 - 데이터 형태

- 넓은 형태 데이터, 짧고 굵은, 가장 기본적인 데이터 형태(Pivot 결과물)
- 컬럼 간 계산이 용이하나, 축 설정 및 데이터 제한이 어려움
- 긴 형태 데이터, 길고 얇은 데이터 형태, 기준 칼럼이 존재(Unpivot 결과물)
- 차트 구성 및 데이터 제한이 용이하나, 데이터를 한눈에 보기 어렵다.
- 피벗 : 긴 자료를 넓은 자료로 변환
- 언피벗 : 넓은 자료를 긴 자료로 변환

● 데이터 핸들링 - 언 피벗(Unpivot)

- 항목들이 한 열에 있지 않고, 각각의 컬럼으로 존재하는 경우
 - 넓은 데이터에서 긴 데이터로 변환, 범주가 나뉜 열 이름은 기본적으로 사용한다. Category가 했던 결과이다.
- 데이터 캔버스-중간+클릭-Transform data-Unpivot-Insert-바깥 데이터는 아래칸, 나머지는 모두 위칸에 옮김

● 데이터 핸들링 - 피벗(pivot)

- 긴 데이터를 넓은 데이터로 변환
- 데이터 캔버스-중간+클릭-Transform data-pivot-Insert-순서대로 값 대입-이름 재정의 필요

● 참고

- 한 데이터에 대해 두 가지 차트를 만들려고 하면 일단 우클릭으로 복사를 하고, 우클릭으로 차트를 바꾸는 방법 사용.
- 꼭 데이터 저장할 때 3번째 꺾로 저장해서 제출하기

----- 기말고사 -----

● 맵 차트

- Spotfire에서는 기본적으로 2개의 레이어(층)을 제공한다.
- 레이어(층)이란 : 여러 계층을 사용하여 데이터를 지리적 컨텍스트에 표현한 차트

- 좌표 참조 시스템(coordinate reference system)
- 동일한 데이터를 각기 다른 참조 시스템을 사용한 예, 좌표 참조 시스템은 맵 차트의 Appearance 탭과 데이터 계층의 setting 창에서 각각 설정 가능, 각기 다른 좌표 참조 시스템 사용시 올바른 위치 지정을 위해 데이터 계층의 좌표를 다시 계산
- Shape 파일(.shp) : ESPI 사에서 제작한 일종의 지도 형상 파일. 속성, 좌표, 도형 등을 포함하고 있는 링크 형식파일로 기본적으로 속성정보와 도형정보를 가지고 있으며, 좌표나 공간 인덱스 정보를 통해 2D, 3D 형식으로 나타낼 수 있는 파일을 일컫는 말이다.
- 맵 차트 : 사용자가 각각의 레이어를 쌓아서 원하는 방식을 표시하는 방식으로 6개까지 쌓을 수 있다.
- 레이어는 크게 2개로 구분한다. 데이터 계층, 참조 계층
- 데이터 계층 : Data Layer : 데이터 계층은 정보를 전달할 핵심적인 내용의 데이터를 포함,
- 표식계층 : 배경이 될 레이어가 필요하며, 좌표 값을 포함한 데이터를 마커로 표시하는 계층이다.
- 기능 계층 : Shape 파일을 통해 원하는 구역, 모양, 형상별로 맵 구획을 나누고 정보를 표현하는 데이터 계층의 하나
- 참조 계층(Reference Layer) : 데이터 계층의 정보를 잘 이해할 수 있도록 보조적인 정보를 제공, 참조 계층만으로는 어떠한 정보도 전달할 수 없음.
- 맵 계층 : 기본적으로 온라인 상태에서 사용가능, 데이터 층없이 직접 상호작용 불가, Spotfire에서 맵 계층을 디폴트로 사용
- WMS 층 : 인터넷에서 다양한 GIS를 제공하기위해, 웹 표준인 웹 맵 서비스에서 검색이 된다.
- TMS 층 : 주로 사용하는 방법이며, 웹 상에서 타일 형태의 조각들로 구성되어 제공하는 지도
- 이미지 계층 : 사용자가 원하는 그림을 만들 수 있으며, x, y를 직접 만들어서 붙이는 작업이 필요하다. 오래걸린다.

[속성]-[Leyer]-Marker Layer...클릭-[setting]-[Positioning]-Coordinate columns클릭-위도, 경도 설정(한글, 영어일 때는 직접 설정, 근데 영어는 자동으로 설정이 될 확률이 높다.)-집계방법은 꼭 None으로 설정한다.

- 또 다른 층을 만들기 위해서는

[속성]-[Leyer]-[Add]-[setting]-여기서 새롭게 세팅

- 마커 크기 조정하는 방법

[속성]-[Leyer]-Marker Layer...클릭-[setting]-[Size]-Marker size

- 원하는 값을 색상에 표시하여 값을 표시한다.

- 기본적인 레이어는 전세계를 표시하는 지도이며, 우리가 표시하는 대부분의 지도는 아파트 동호수까지 나와야 한다. 따라서 다른 지도를 다른 레이어로 사용해야 한다. 지금 사용하는 것은 우리나라 사람들이 우리가 쓰려고 만든 TMS 레이어를 가져온다.

[속성]-[Layer]-[add]-[Map Layer]-[TMS Layer]-url에 해당 값을 입력-update-

기본 : <http://xdworld.vworld.kr:8080/2d/Base/service/{z}/{x}/{y}.png>

그레이 : <http://xdworld.vworld.kr:8080/2d/gray/service/{z}/{x}/{y}.png>

블랙 : <http://xdworld.vworld.kr:8080/2d/midnight/service/{z}/{x}/{y}.png>

구글 일반 맵 : <https://mts1.google.com/vt/lyrs=m&hl=ko&x={x}&y={y}&z={z}>

구글 일반 맵 + 등고선 : <https://mts1.google.com/vt/lyrs=p&hl=ko&x={x}&y={y}&z={z}>

구글 일반 맵 + 등고선 + 길우선 : <https://mts1.google.com/vt/lyrs=r&hl=ko&x={x}&y={y}&z={z}>

구글 위성사진 : <https://mts1.google.com/vt/lyrs=s&hl=ko&x={x}&y={y}&z={z}>

구글 일반 맵 + 등고선 : <https://mts1.google.com/vt/lyrs=t&hl=ko&x={x}&y={y}&z={z}>

구글 위성사진 + 한글지명/건물명 : <https://mts1.google.com/vt/lyrs=s&hl=ko&x={x}&y={y}&z={z}>

- 레이어 선택, 현재 만든 층을 조정하는 방법

왼쪽 위 DB표현과 비슷한 아이콘 클릭-원하는 층을 클릭하여 쌓기, 순서대로 쌓이기 때문에 순서 조정을 하며 밑에서부터 투명하게 쌓아야 한다.

- 불러온 Shape파일이 한글로 인한 깨짐 현상을 보인다면

해당 cell 우클릭-[Replace Value]-바꿀값으로 지정-Select key columns선택-Apply-작업을 반복하면서 값을 다시 지정

- 두 개의 파일을 불러와서, 컬럼을 연동한 다음, 지도에 표기하는 방법

두 개의 파일을 불러온다.(데이터 레이어, 참조 레이어)-[Data]-[Data Table Properties]-[Column Matches]-new-기준 행을 지

정-Apply-맵 차트 열기-Feature by에 Shape 파일 기준 열을 설정-color by에 데이터 레이어 데이터를 열어서 표시할 data를 표시(오른쪽 상단의 데이터 바꾸는 것을 유의)

- 두 데이터의 연동할 열이 잘 보이지 않는 경우

두데이터를 불러와서-열을 자르거나, [Data]-[Add calculated column]을 사용하여 기준열을 생성-[Data]-[Data Table Properties]-[Column Matches]-new-기준 행을 지정-Apply-맵 차트 열기-Feature by에 Shape 파일 기준 열을 설정-color by에 데이터 레이어 데이터를 열어서 표시할 data를 표시(오른쪽 상단의 데이터 바꾸는 것을 유의)

- 표시되는 지역에 레이블 형식 바꾸기

[Data]-[Column Properties]-표시되는 데이터와 컬럼 클릭-Formatting-Category-Custom-Format string 작성-맨 우측 하단 Apply 클릭-OK클릭

● 라인 & 곡선

- 시각화에 참조선이나 통계저긴 수치를 표시

- 정성적인 부분을 정량적으로 확인가능

- 데이터 타입별 Vertical Line 활성화.비활성화 차이 존재(x축이 string이라서 그렇다.)

● 추세선 그리기

[속성]-[Line&Curves]-[Straight line fit : straight line] 클릭-[Label and Tooltip]-결정계수 표현

● 수직선과 수평선

- 수직선 추가(Fixed : 고정된 값으로 수직선, Aggregated value : 평균으로 수직선)

[속성]-[Line&Curves]-[Add]-[Horizontal Line]-[straight line]-알맞은 기준 설정

- 수직선에 값 표시하기, 선 종류, 굵기 설정

[속성]-[Line&Curves]-Visible lines and curves에서 해당 선 선택-[Label and Tooltip]-해당 옵션 선택

- 레이블 형식 변경

[속성]-[Line&Curves]-Visible lines and curves에서 해당 선 선택-[Label and Tooltip]-[Format]-형식 변경

- 표준편차*(+- 1)로 표시하기

[속성]-[Line&Curves]-[Add]-[Horizontal Line]-[Average and +- 1 standard Deviations]

- 표준편차*(+- 3)로 표시하기

[속성]-[Line&Curves]-[Add]-[Horizontal Line]-[Average and +- 1 standard Deviations]-해당 선을 클릭 후-[Edit]-[Edit]-수식에서 1을 3으로 변경-OK-OK

- 수평선도 위에 수직선과 똑같이 시행하면 된다.

● 추세선 그리기

- 5차 함수까지 표현 할 수 있는 Polynomial Curve Fit 사용

[속성]-[Line&Curves]-[Add]-[Polynomial Curve Fit]-원하는 차수 선택-OK

- 자세한 레이블 설정은 Label and Tooltip에서 설정

● 가우스 추세선 사용하기, Gaussian Curve Fit 사용

- E값이 높으면 평균이 오른쪽으로 치우쳐진다(값이 커진다)라고 볼 수 있다.

[속성]-[Line&Curves]-[Add]-[Gaussian Curve Fit]-자동값으로 설정한다-[Label and Tooltip]-E값 출력

- Trellis by나 열의 순서가 잘못되어 있는 경우 조정하는 법

[Data]-해당하는 컬럼선택-[Sort Order]-Custom sort order-Configure-열 순서 조정-OK-Apply

- y축에 눈금을 표시하기 위해서 Y축을 클릭하여 '쇼더 가이드라인'을 클릭한다.

● 예측선 그리기, Forecast-Holt-Winters

- 모수의 개수의 1/10까지 낮은 오차로 예측 가능하다.

[속성]-[Line&Curves]-[Add]-[Forecast-Holt-Winters]-Time point ahead: 몇 개를 예측할 것인지 표시-OK

- 자동으로 해서 적합하는 것이 좋지만 사용자 지정을 한다고 하면
- 1. Level(알파) 수준 : 무조건 해야되는 옵션으로 0에서 1사이 값을 가지고, 0은 과거, 1은 현재에 가까운 데이터에 가중치를 높게 부여한다는 적합, 예) 0.75이면 최근 데이터 25%에 비중을 더 두고 예측을 한다
- 2. Trend(베타) 추세 : 장기적으로 값이 증가, 감소 등 추세를 보일 때 사용, 기울기라고도 한다. 0에서 1사이 값
- 3. Seasonal(감마) 계절성 : 값이 반복성 주기성을 보일 때 사용, 0에서 1사이의 값
- 시계열의 주시성이 강조되면 Additive,
- 시계열의 진폭이 점점 증가감소하면 Multiplicative
- 데이터를 한정적으로 선택하여 차트를 보고싶으면 범위를 줄여서 보고 싶은 Y, X축 클릭-Show Zoom Slider

● 데이터 상관성 분석

- 열 데이터의 형식을 알아보는 방법
[Data]-[Column properties]
- 크로스 테이블을 같이 사용하면 좋다. 데이터를 요약해주는 테이블이기 때문이다.
- 해당 데이터의 추세선 기울기 그리기
산점도 차트-[속성]-[Line&Curve]-[Add]-[Straight Line Fit]-Costom으로 해서 이름 설정-[Label and Tooltip]-(b)값 표시
- 해당 데이터의 추세선 그리고 각종 데이터 표시하기
산점도 차트-[속성]-[Line&Curve]-[Add]-[Straight Line Fit]-Costom으로 해서 이름 설정-[Label and Tooltip]-모든 값에 클릭
- Treills를 사용하여 범주형인 요인과의 비교가 가능하다
- 데이터 상관분석 : 선형 회귀
[Tools]-[Data relationships...]-이름설정-옵션으로 Linear Regression 선택-y축에 종속변수, x에 독립변수 대입-확인
- P값 소수점 포맷 설정
[Data]-[Column properties]-p값 선택-Format-숫자-소수점 6자리 설정-확인
- 정령이 안 되어있으면, P-value와 결정계수 정렬을 다시 해준다. 아래 선형 표시 시각화에 Trellis 사용 가능
- P-value가 0.05보다 작으면 x컬럼은 y컬럼에 유의미한 영향을 준다.
- P-value 가설이 맞을 확률로 0.05 즉 5%보다 작으면 가설을 기각한다.
- 가설 $b=0$
- 상관계수(R) : $-1 < R < 1$, 상관계수는 절댓값이 1에 가까울수록 x와 y가 큰 관계성을 가진다.
- 결정계수(R^2) : $0 < R^2 < 1$, 1에 가까울수록 x와 y가 큰 관계성을 가진다.
- 데이터 상관분석 : 스피어만 상관분석 Spearman-R ,
- 이 방법을 쓰면 그래프가 하나 더 생성되는데 이 그래프에서 RankReal는 순위를 정해주는 방법으로 행이 30개 미만인 적은 데이터를 사용할 때 사용.
- anova 분산분석]
- 상자그림이 나온다.
- 상자그림 설정하는 방법
[설정]-[Reference Points]-표시하고자 하는 값들을 클릭-색상과 마크 표시를 설정
- 상관성 원 교차 시각화 표시하기, 원이 멀어질수록 집단간 차이가 있다라는 것을 뜻함, 원의 수는 x에 해당하는 범주형 변수의 집단 수를 뜻함
[속성]-[Appearance]-[Show comparison circles]
- P-Value값이 0.05보다 작으면 평균의 차이가 유의미하다(있다.)

- 귀무가설 : 그룹 간 평균의 차이가 없다
- .P-Value값이 0.05보다 작으면 귀무가설 기각, 즉 차이가 있다고 볼 수 있다.

- Kruskal-Wallis
- 그룹별 행 수가 5이하일 때 사용하는 변량분석

- 카이제곱검정
- 문자열과 문자열의 그룹별 비율 차이 검토할 때 사용.
- P값이 0.05 보다 작으면 그룹별 비율의 차이가 있다고 본다.
-

● 라인 유사성

- y축에 멀티를 거는 방법
y축을 우클릭한다.- Multiple Scales을 설정
- 데이터를 요약하기 위해 옆에다가 크로스 테이블을 표시하는 것도 좋은 방법이다.

- 라인 유사성 설정

유사성을 검사할 라인 그래프를 마킹-[Tools]-[Line Similarity]-이름 설정-확인-필터일에서 새롭게 생긴 similar rank를 트릴로 추가한다.

- 라인 유사성을 가지고 좀 더 자세하게 표현하는 방법

[속성]-[Trellis]-[Panels]-[Manual layout]-격자 순서 확인- Split by 로 Similar rank, Similar, 색상으로 표현하고 있는 요인 대입(그렇게 된다면 트릴 각 상자 위에, 비슷한 순위, 상관계수, 분류된 이름을 표시할 수 있다.)

● K 평균 군집 분석

- k(나눌 집단의 수)를 정하여 데이터를 여러군집으로 나누는 방법
[Tools]-[K-means clustering...]-k에 분류할 집단의 수 입력-확인

- 크로스테이블을 이용하여 같이 활용해도 좋음. 범주형데이터에 들어가는 이산데이터들을 표시하는 방법

크로스테이블 생성-세로에 군집분석한 컬럼 대입-위쪽은 삭제-아래는 점주형 데이터를 클릭-UniqueConcatenate로 생성