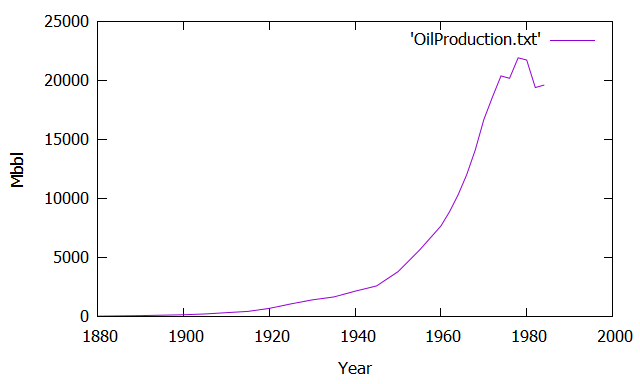
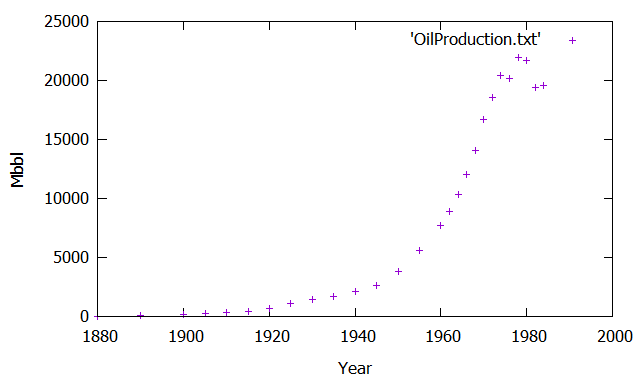
1. Draw graphs using GNU plot

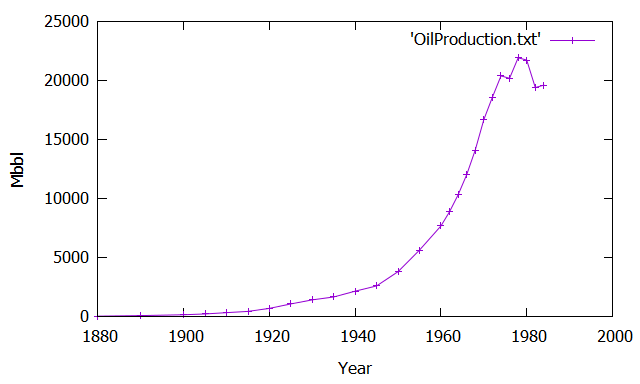
* Line



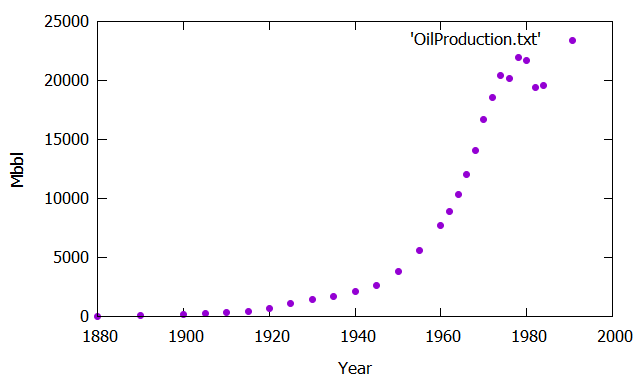
* Point (scatterplot)



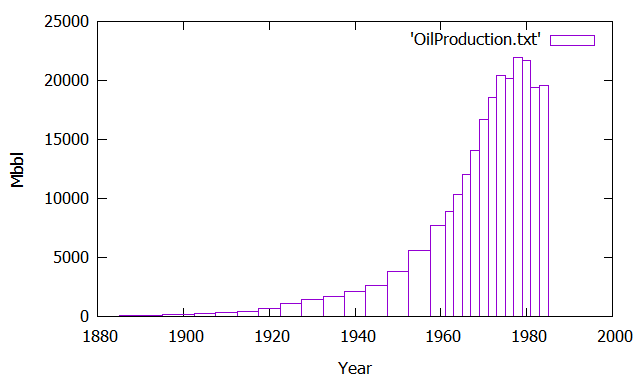
- Line + point



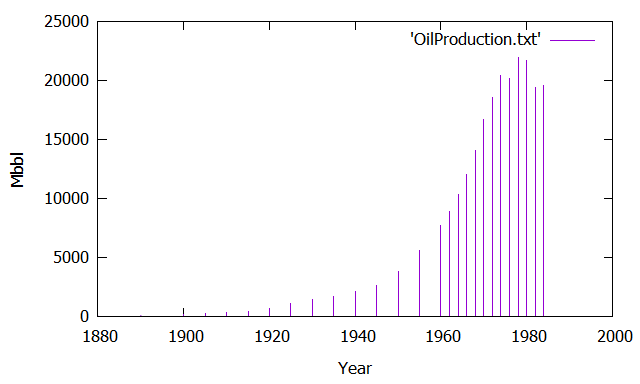
* Dots



* Boxes



* impulses



1. Data transform
2. x: Year, y: Mbbl 이라고 할 때, 다음과 같은 exponential 모델을 생성하였다.

Parameter a와 b를 찾기 위해, 다음과 같은 계산 과정을 거쳤다.

Let ,

then (linear)

위와 같은 변환을 통해, non-linear 모델을 linear model 로 변환하였고,

변환된 linear model의 parameter인 , 을 찾고자 한다.

모델의 값과, 실제 값의 차이를 나타내는 SSE를 다음과 같이 정의하자.

=

SSE를 최소화하는 , 값을 찾기 위해, 다음과 같은 미분과정을 거친다.

=

= = 0

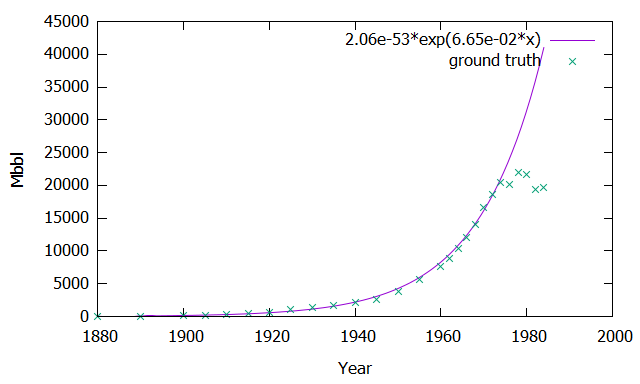
계산을 통해 아래와 다음과 같은 식을 얻어낼 수 있다.

(값의 평균, : x 값의 평균)

* , b

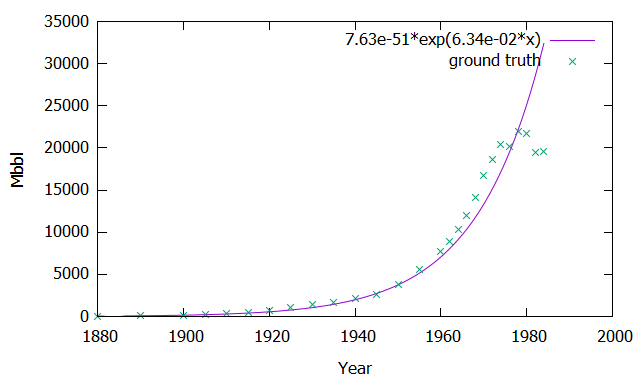
위 식에 값들을 대입하여 계산하기 전, 두 가지 모델을 생각하여 보았다.

1. 1973년 이전의 데이터만을 이용한 모델 – (모델 1)
   * + - 1973년 전의 데이터는 정치적인 영향이 없다는 것을 고려하여 data의 패턴을 더 잘 표현하는 데이터라고 판단하였다.
       - 1917년 이후의 데이터는 고려하지 않았다.
       - Parameter: -121.3137, 0.0665 (parameters of linear model)

=> a 2.06e-53, b 6.65e-02 (parameters of exponential model)

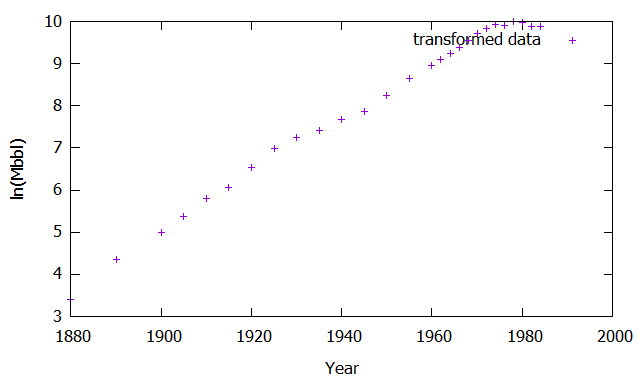
[그림 1]

1. 전체 데이터를 이용한 모델 – (모델 2)
   * + - 주어진 전체 데이터를 반영하였다.
       - Parameter: -115.4003, 0.0634 (parameters of linear model)

=> a 7.63e-51, b 6.34e-02 (parameters of exponential mode

[그림 2]

1. Mbbl의 값에 ln을 취해 log 변환을 수행하였다.
2. Scatter plot of log transformed data



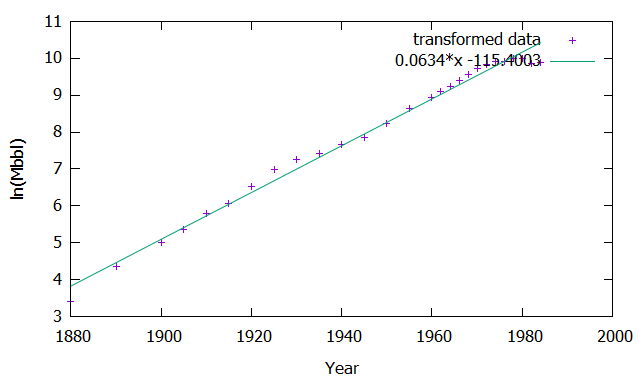
1. Provide your own analysis and discussion about the oil production.

* 1973년 이후 정치적인 혼란 또는 다른 요소의 영향이 없었다면, [그림 1] 과 같이 원유 생산량이 급격하게 증가했을 것이라고 예측할 수 있다.
* 지수적으로 증가 혹은 감소하는 데이터를 모델링 할 때,

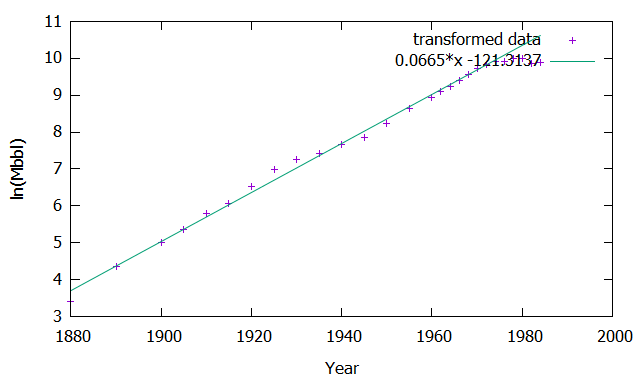
데이터에 자연로그를 취하면 선형 모델로 회귀분석이 가능함을 알 수 있었다.

Mbbl에 자연로그를 취하여 변형된 데이터는 위에서 구한 파라미터들을 이용하여 아래와 같이 단순한 선형 모델로 표현할 수 있다.

-121.3137, 0.0665 (parameters of model 1)



-115.4003, 0.0634 (parameters of model 2)

****