

Two hexagons are positioned in the top right corner. The one in the foreground is light gray, and the one behind it is a vibrant blue.

제조 빅데이터 **전문가** 과정

빅데이터 기반의 생산성 효율

(주)에스투비즈
대표컨설턴트
이 이 백

Content



03 제조 빅데이터 분석

04 품질 최적화

05 설비 예지보전 및 생산 최적화

06 제조산업에서의 딥러닝 활용

02

품질 최적화

품질 최적화

Process

Quality정보 수집

검사/측정

Exploration

Analytics

사후 조치

- 4M
 - 자재(material)
 - 설비(machine)
 - 작업(method)
 - 작업자(man)
- WIP History
- Production History
- FDC Parameter

- CTQ Parameters
- 양품과 불량 식별
- 자주검사
- Vision 측정 검사

- Chart
- Pareto
- 공정 능력
- SPC
- Real Time Monitoring

- Root Cause Analytic
 - 비교자원간 유의차 분석
- Prediction
- Forecasting

- Recipe 조정을 위한 DoE
- 지능화 프로젝트 추진
- Run To Run

품질 최적화



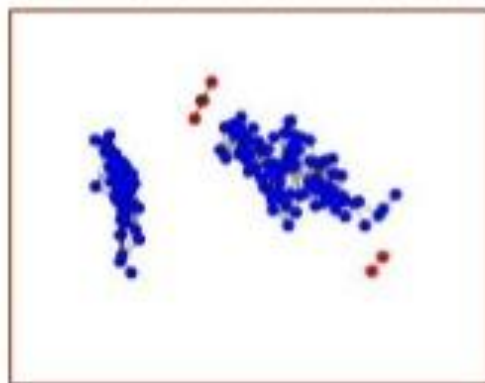
Quality

- Cost of Quality
- Quality : 설계, 부품, 제조, 설비의 요구기능과 성능의 완전성
- 품질관리 활동
 - 불량원인의 조기검출
 - 고객사에 위음성(false negative) 제품의 반출 예방
 - 고객 이슈의 공정 추적 및 원인 규명
- Analytics
 - Abnormal Detection – SPC, Timeserise, PCA, Autoencoder
 - Root Cause Analysis : Decision Tree, Random Forest, Autoencoder
 - Image 분석 – CNN
 - Clustering : PCA, Kmeans, TSNE
 - Prediction : Yield, Demand Forecasting
 - ML - Support Vector Machine, Random Forest
 - DL – MLP, CNN, RNN

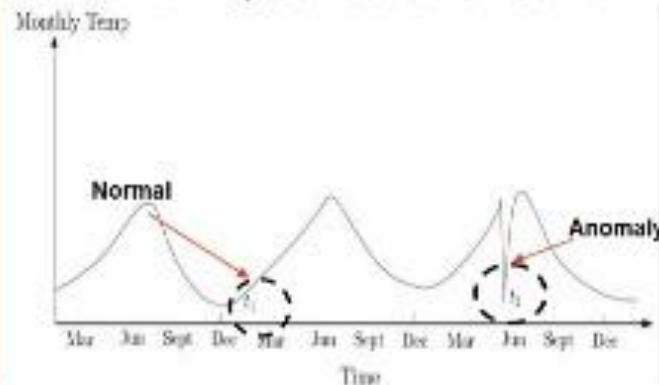
Anomaly Detection

TYPES OF ANOMALIES

- Anomalies can be classified as Point , Collective or Contextual .



- Point Anomaly**
- If an **individual data** instance can be considered as anomalous with respect to the rest of the data (e.g. purchase with large transaction value)



- Contextual Anomaly**
- If a data instance is anomalous in a specific **context**, but not otherwise (anomaly if occur at certain time or certain region. e.g. large drop at middle of month in consumption)

Anomalous Subsequence

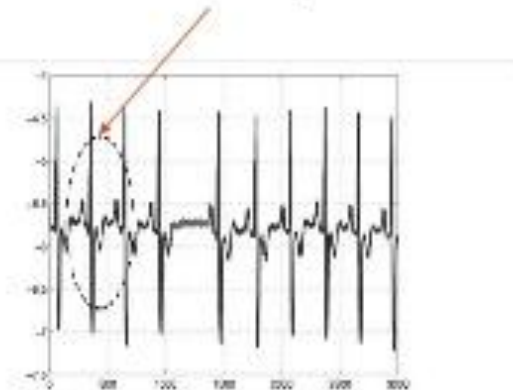


Fig. 4. Collectively anomaly corresponding to an Anomalous Subsequence in the case of understanding the output.

- Collective Anomaly**
- If a collection of related data instances is anomalous with **respect to the entire data set**, but not individual values (e.g. breaking rhythm in ECG)



Anomaly Detection

단변량

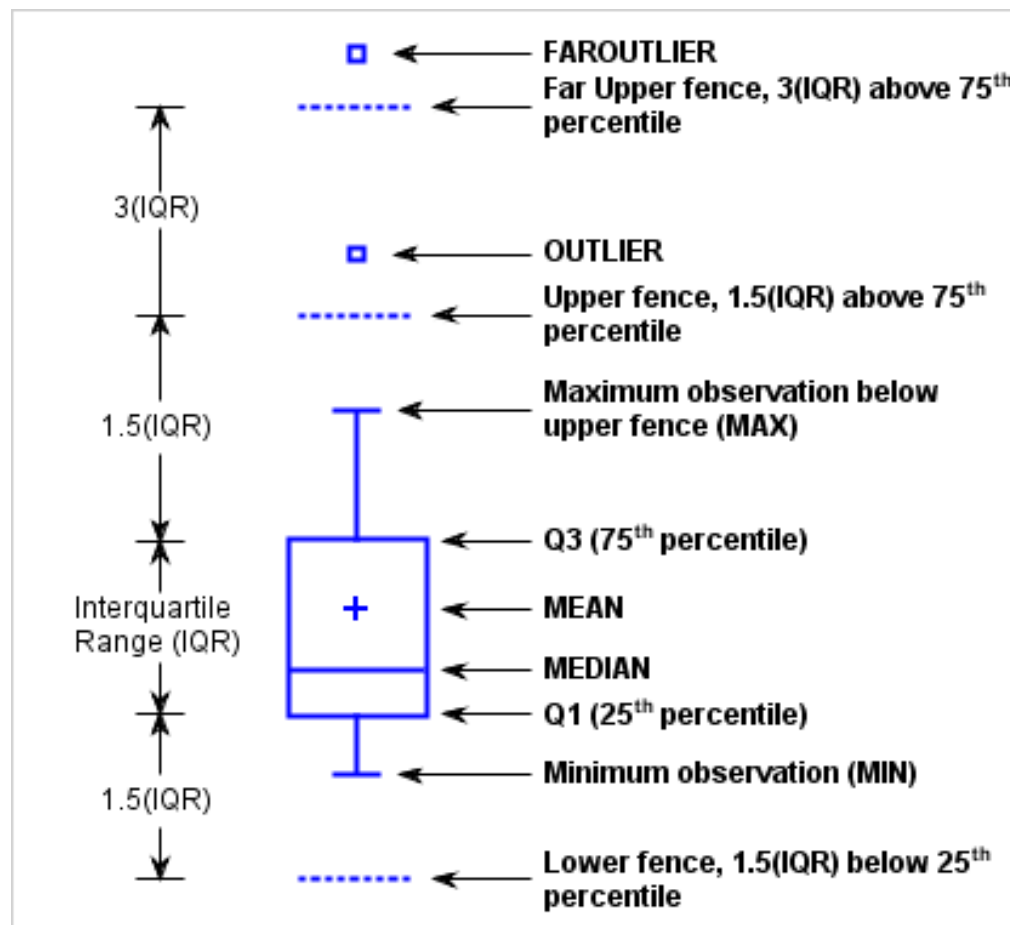
IQR

Boxplot = Box Whisker Plot

$$IQR = 3Q - 1Q$$

이상치 기준

- 상한 = $3Q + IQR \cdot 1.5$
- 하한 = $1Q - IQR \cdot 1.5$

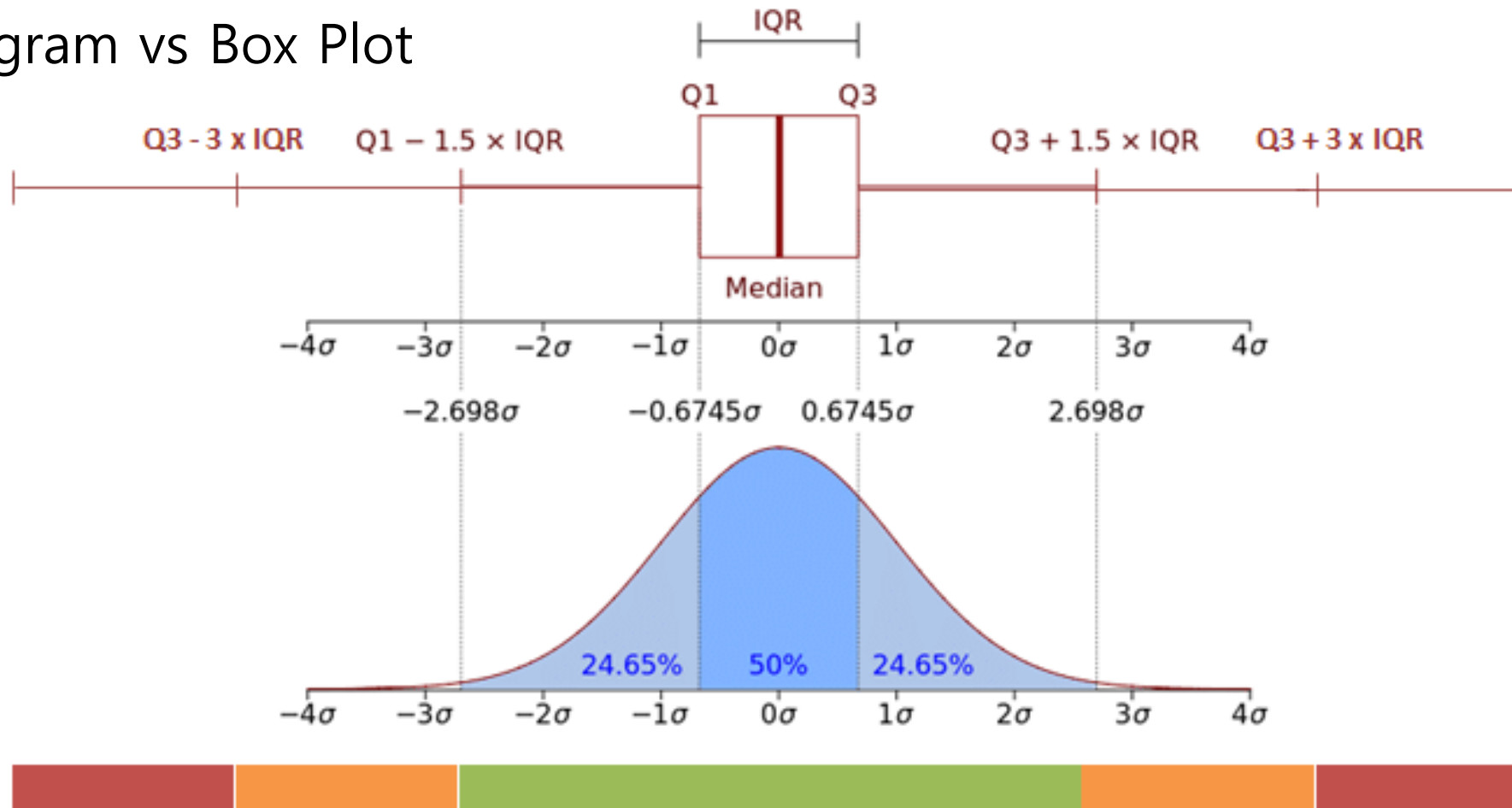


Anomaly Detection

단변량

IQR

■ Histogram vs Box Plot

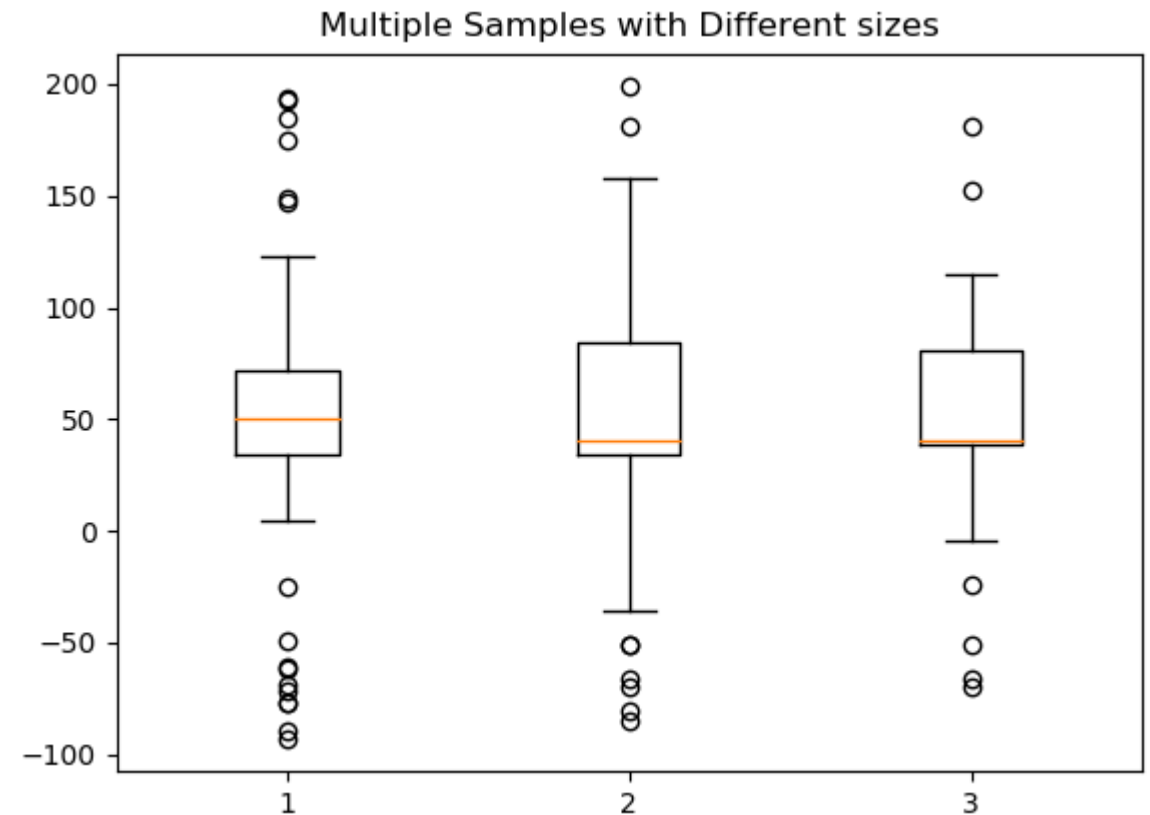
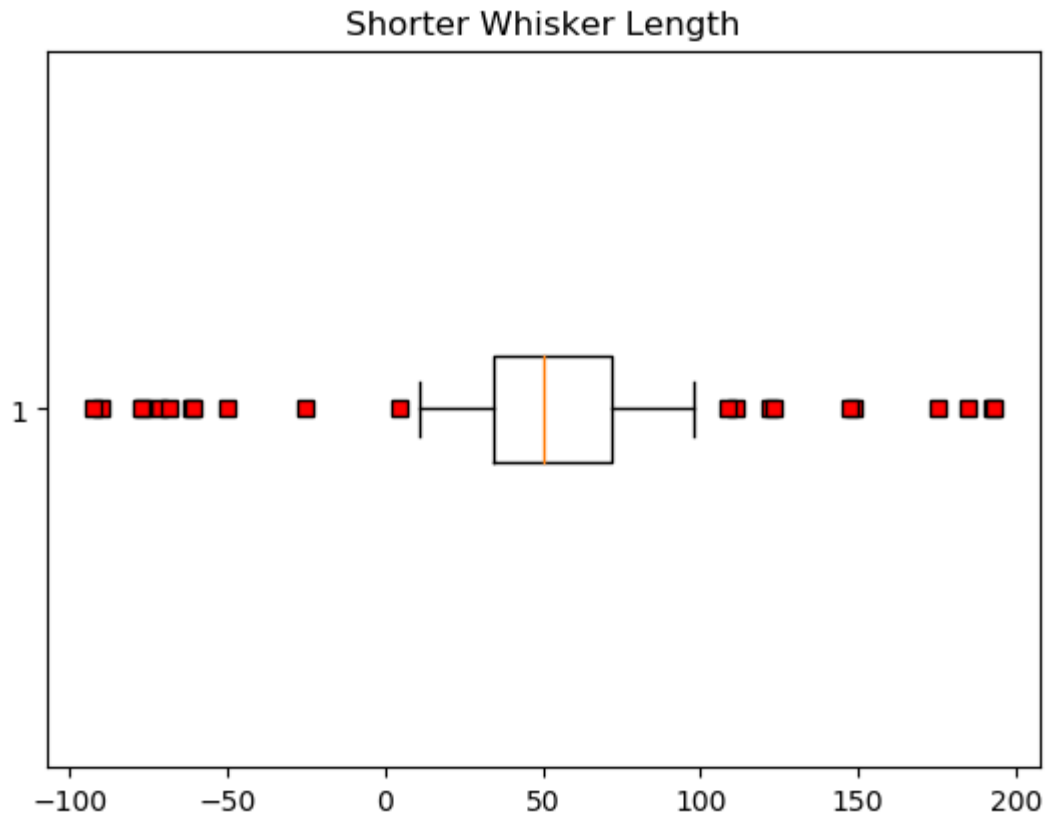


Anomaly Detection

단변량

IQR

■ Box Plot/IQR을 이용한 Parameter의 이상치 검토



Anomaly Detection

단변량

SPC

- 시간의 흐름에 따른 공정의 변화 식별
- 공정이 얼마나 안정적인지 불안정적인지 판단
 - Current \rightleftharpoons History
- 기준 data : Historical data
- 목적에 따른 Control Chart 유형
 - X-bar chart
 - R chart
 - S chart
 - P chart
 - NP chart
 - C chart
 - U chart

SPC 유형	내용
Xbar	<ul style="list-style-type: none"> ▪ Sample Means ▪ Parameter의 평균값
R	<ul style="list-style-type: none"> ▪ Parameter평균의 변동 범위
S	<ul style="list-style-type: none"> ▪ Parameter의 표준편차
P	<ul style="list-style-type: none"> ▪ Portion of Nonconfirming ▪ 불량율
NP	<ul style="list-style-type: none"> ▪ Number of nonconfirming Units ▪ 불량수
C	<ul style="list-style-type: none"> ▪ Number of defectives per unit ▪ Defect 수
u	<ul style="list-style-type: none"> ▪ Average number of defectives per unit ▪ 단위당 평균Defect 수

04. 품질 최적화

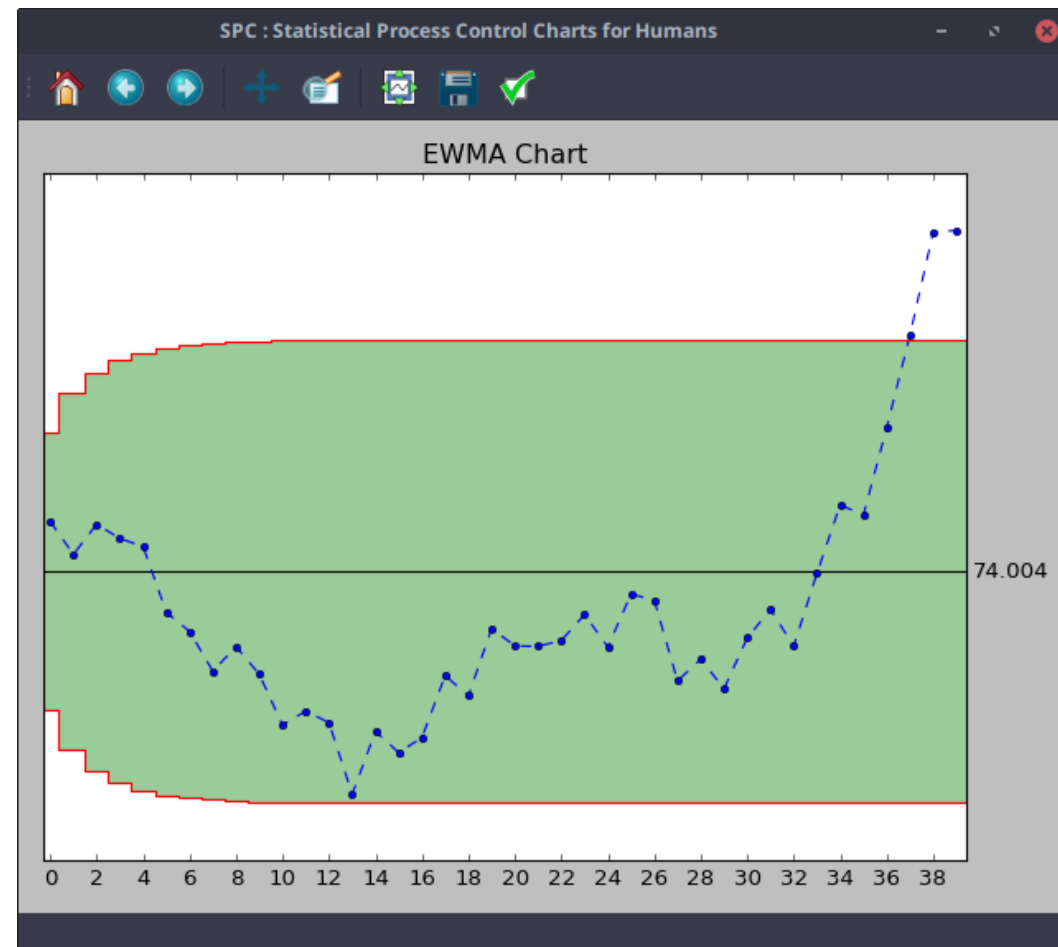
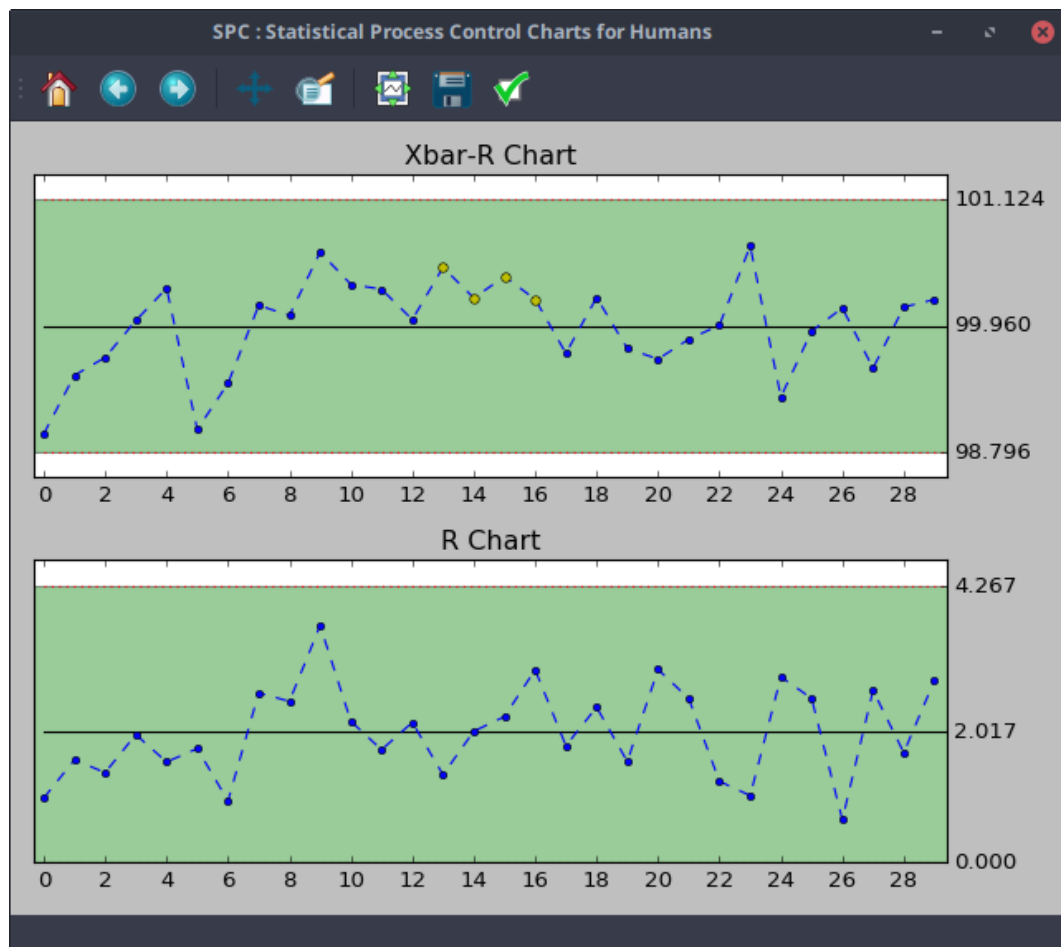
Anomaly Detection

11

Anomaly Detection

단변량

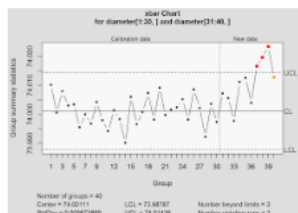
SPC



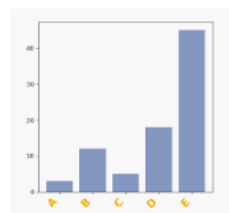
Anomaly Detection

단변량

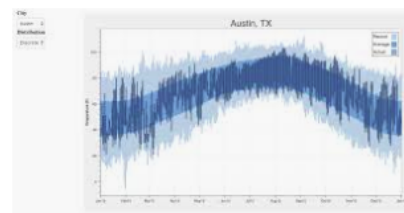
SPC



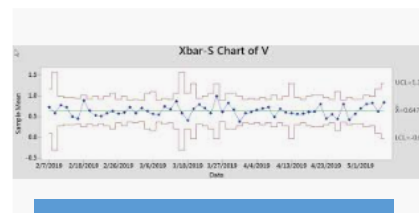
Control Charts in R ...
datascienceplus.com



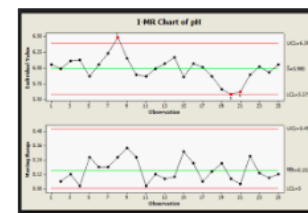
All Charts - The Python Gr...
python-graph-gallery.com



Python Data Visualization Libraries ...
mode.com



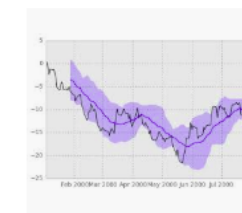
Control Charts in Six Sigma ...
greycampus.com



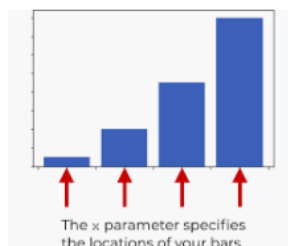
Control Charts in Six Sigma ...
greycampus.com



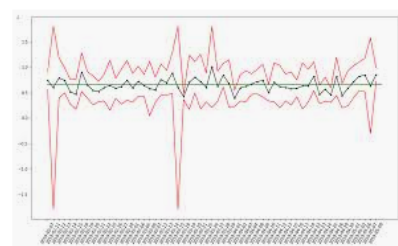
Easier chart creation in Python ...
labs.spotify.com



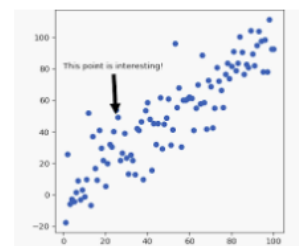
Plotting - pandas 0.15.0 ...
pandas.pydata.org



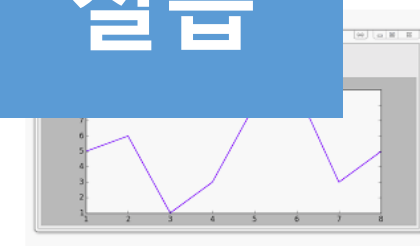
How to make a matplotlib bar chart ...
sharpsightlabs.com



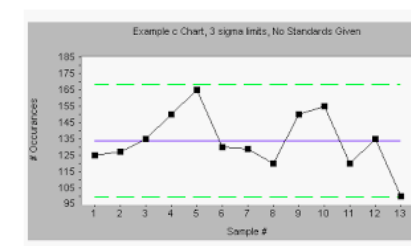
Minitab style control chart ...
stackoverflow.com



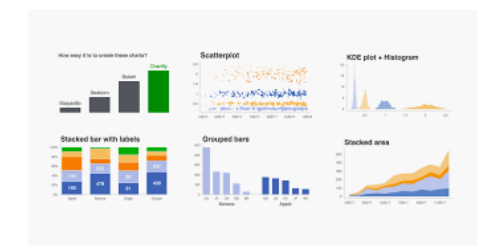
All Charts - The Python Graph Gallery
python-graph-gallery.com



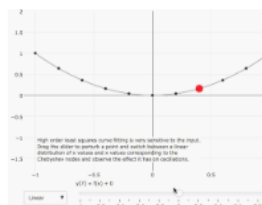
Python Programming Tutorials
pythonprogramming.net



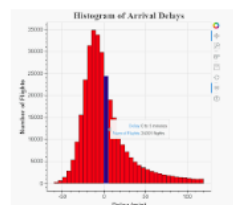
Control Charts in Six Sigma ...
greycampus.com



Easier chart creation in Python ...
labs.spotify.com



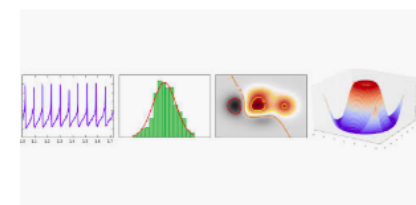
15 Python and R Charts with In...
moderndata.plot.ly



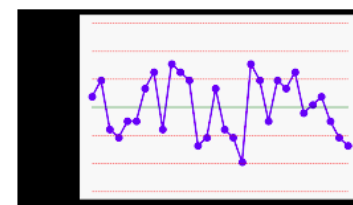
Data Visualization with Bok...
towardsdatascience.com



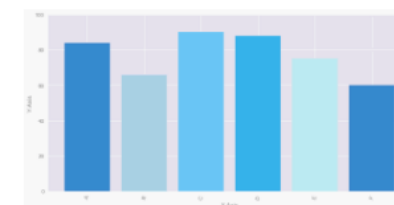
Matplotlib - The Python Graph G...
python-graph-gallery.com



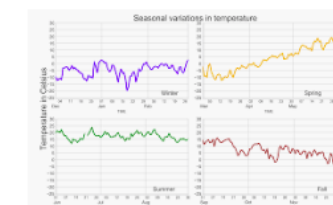
matplotlib: python plotting ...
omz-software.com



Control Chart with Scripted Rule Checks
joshuaschultz.com



Python Plotting Basics - Towards Data ...
towardsdatascience.com



Advanced plotting with Pandas - Geo ...
geo-python.github.io

실습

Anomaly Detection

다변량

비지도 학습

PCA

- 고차원 데이터를 정보의 손실을 최소화하며 저차원으로 변환하는 기법
- 데이터 차원(=변수)의 선형 결합을 통해 새로운 특징 차원(=주성분변수) 즉 주성분을 산출
 - 상관계수 또는 분산-공분산을 이용하여 원래변수를 선형결합하여 주성분 변수의 고유치(Eigen Value)와 고유벡터(Eigen Vector)를 산출
 - 결과적으로 주성분 변수간에는 상관성이 최대한 제거되게 함
- 주성분의 양이 많은 소수의 특징 차원이 본래 데이터 차원을 충분히 설명할 수 있으면 원래의 데이터 차원 대신 소수의 특징 차원으로 탐색 및 분석 모델링에 활용할 수 있음
- 탐색 - 주성분 특징변수 2~3개로 시각화
- 군집화 - 특징변수의 값이 비슷할수록 원래변수간 상관성이 큰 경향을 보임
- 모델활용 - 회귀모형의 독립변수로 활용하여 다중공선성 회피
- PCA를 이용한 Anomaly Detection
 - 주성분 2 또는 3개로 산점도 그래프로 시각화하여 군집에서 떨어진 관찰치 탐색
 - 주성분값의 Timeseries 관찰
 - Mahalanobis의 접목

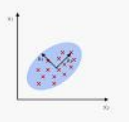
Anomaly Detection

다변량

비지도 학습

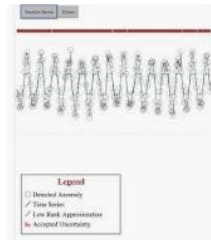
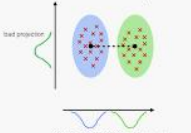
PCA

PCA:
component axes that
maximize the variance

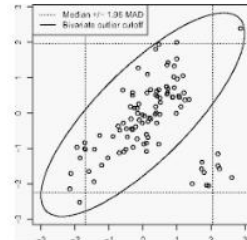


Entry Point Data
sebastianraschka.com

LDA:
maximizing the component
axes for class-separation



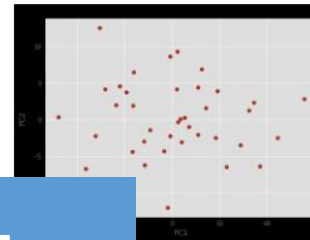
RAD — Outlier Detection o...
techblog.netfix.com



Robust feature selection and r...
semanticscholar.org



Princip...
towards



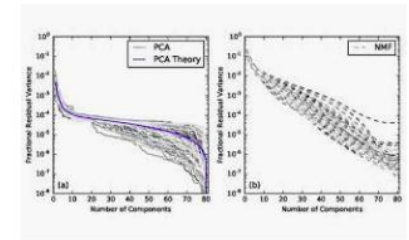
ance with PCA in Python

실습

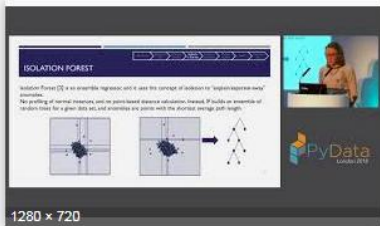
Algorithms on H2O



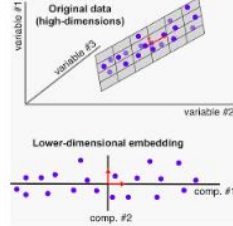
Machine Learning with H2O and Python ...
slideshare.net



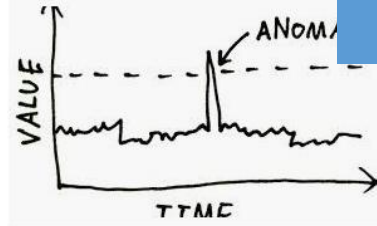
Principal component analysis - Wikipedia
en.wikipedia.org



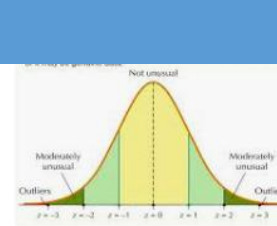
Unsupervised Anomaly Detection with ...
youtube.com



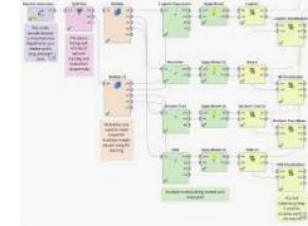
Everything you did and didn't...
alexhwiliams.info



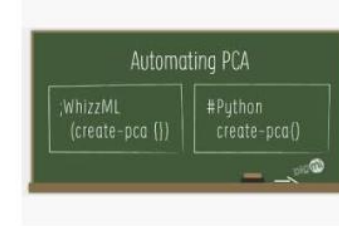
Fundamentals of Machine Learning with ...
d414v1z.com



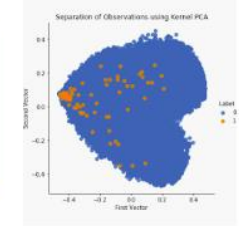
Machine Learning Models Robust to...
heartbest.fritz.ai



Anomaly Detection using Rapidminer an...
medium.com



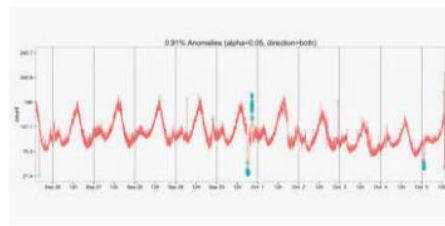
Automating Principal Component Analysis ...
blog.bigml.com



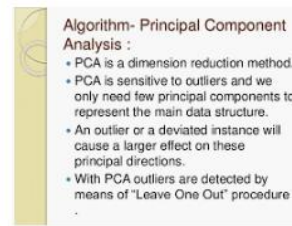
4. Anomaly Detection - Ha...
oreilly.com

Related searches

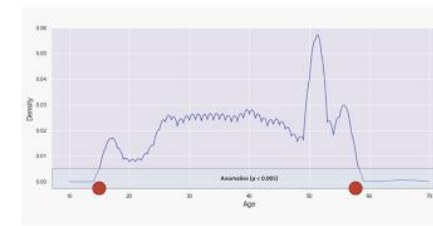
- patient controlled analgesia >
- mnist pca >
- pca image compression >



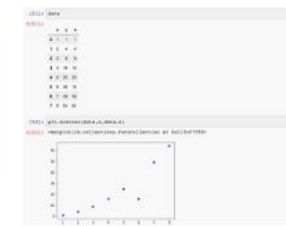
robust anomaly detection in a time series
blog.twitter.com



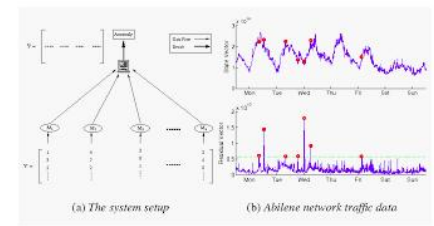
Anomaly Detection Via PCA
slideshare.net



Anomaly detection - Hands-On ...
subscription.packtpub.com



Multivariate Outlier Removal With ...
stackoverflow.com



In-Network PCA and Anomaly Detection ...
semanticscholar.org

Anomaly Detection

다변량

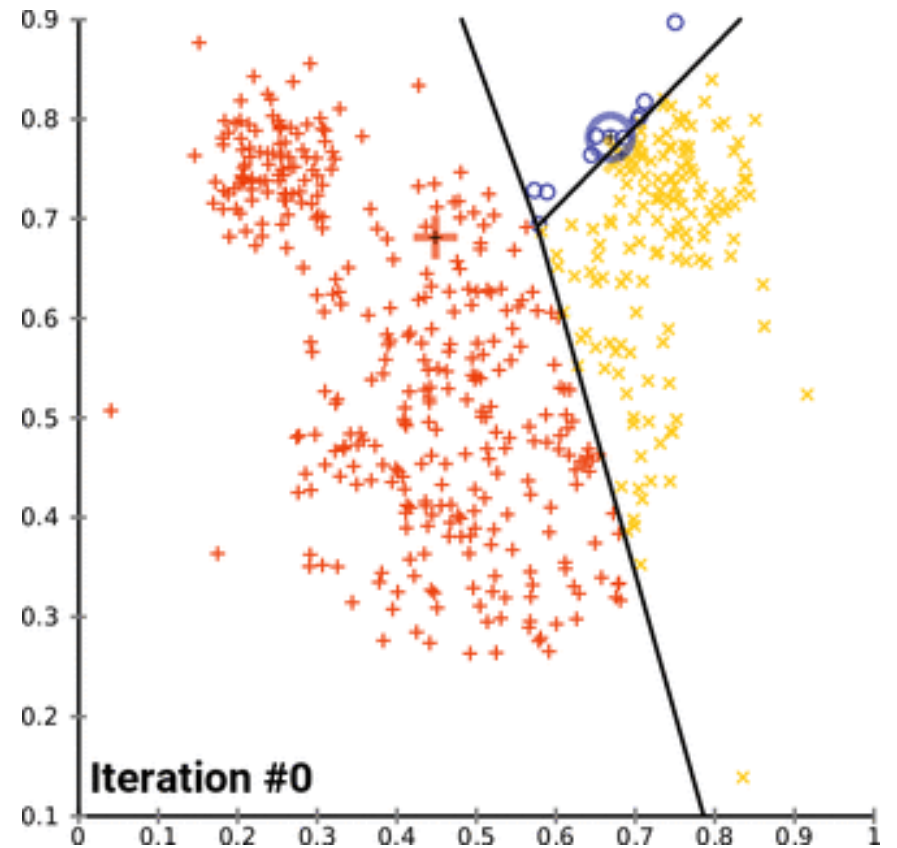
비지도 학습

KMEANS

- 군집분석의 한 종류
- K = 분할할 그룹의 수
 - [1...K] 개별 그룹내의 feature는 유사도가 높음
 - [1...K] 개별 그룹간의 feature는 유사도가 낮음
- Means
 - 각 K 그룹의 무게중심(centroid)
- Formula

$$\arg \min_{\mathbf{S}} \sum_{i=1}^k \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2 = \arg \min_{\mathbf{S}} \sum_{i=1}^k |S_i| \text{Var } S_i$$

- 집단을 분할하는 과정
 - [1...K] 개별 그룹간의 feature는 유사도가 낮음
 - 관찰치를 K개의 집합으로 나눔
 - 각그룹의 임의의 점을 K-Mean value로 선택
 - 관찰치들은 k-Mean을 중심점을 기준으로 Grouping
 - 더 좋은 그룹핑 기회가 있을 때까지 반복해서 k-Means값의 재 조정



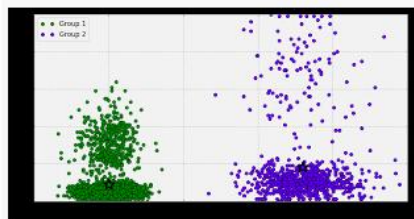
Source : https://en.wikipedia.org/wiki/K-means_clustering

Anomaly Detection

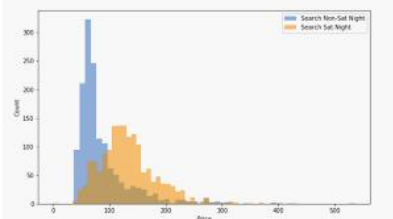
다변량

비지도 학습

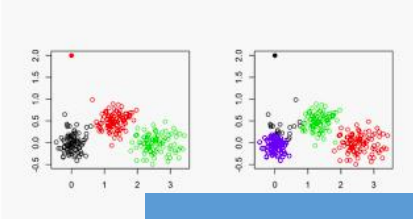
KMEANS



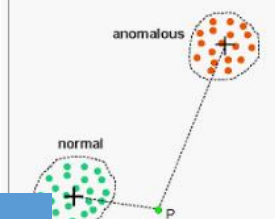
Introduction to K-means Clustering
datascience.com



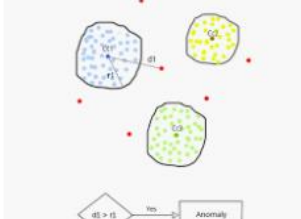
Time Series of Price Anomaly Detection ...
towardsdatascience.com



Outlier detection in
stackoverflow.com



Anomaly Detection Using K-means ...
datascience.com



Anomaly Detection Using K-means ...
ashenweathunga.wordpress.com

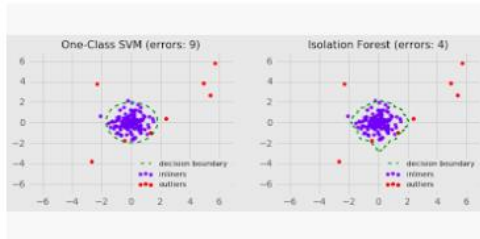
Related searches

k means clustering example

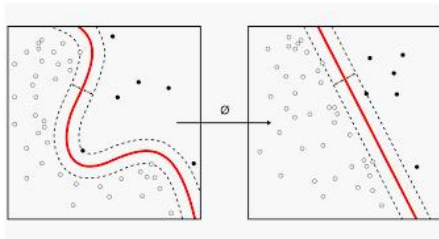
k means clustering gif

k means clustering image segmentation

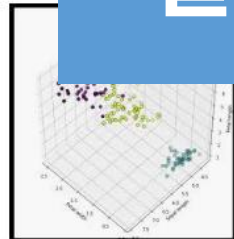
실습



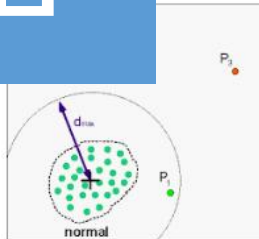
Isolation Forests for Anomaly Detection ...
insidebigdata.com



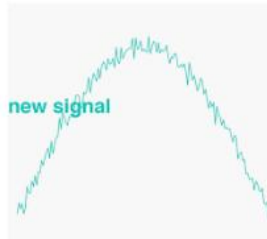
k-means clustering - Wikipedia
en.wikipedia.org



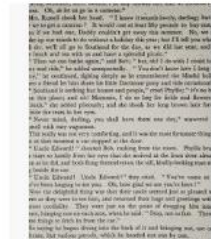
Clustering Based Unsupervised ...
towardsdatascience.com



Traffic Anomaly Detection Using ...
semanticscholar.org



Anomaly Detection Using K-Means ...
anomaly.io



Anomaly detection in Twee...
medium.com



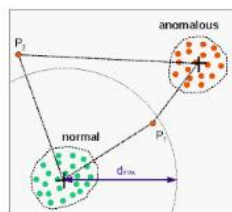
Dis-advantages of k-means clustering ...
inovex.de



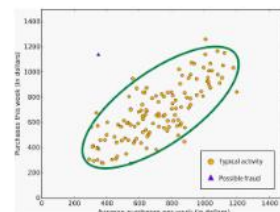
Anomaly Detection with Time Se...
towardsdatascience.com



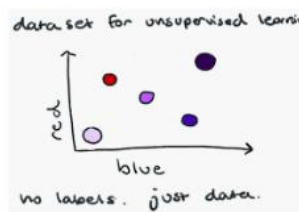
Introduction to anomaly detection in py...
blog.floydhub.com



Traffic Anomaly Detection ...
semanticscholar.org



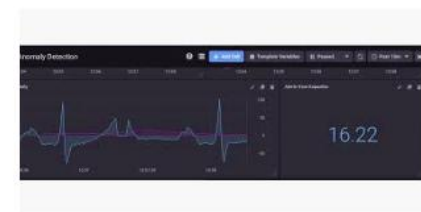
Azure Machine Learning Studio ...
docs.microsoft.com



Why Use K-Means for Time Series Dat...
medium.com

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	ParCh	Ticket	Fare	Cabin	Embarked
0	False	False	False	False	False	False	False	False	False	True	False
1	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	True	False
3	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	True	False

K-Means Clustering with scikit-learn ...
datacamp.com



Why Use K-Means for Time Series Data ...
influxdata.com

Anomaly Detection

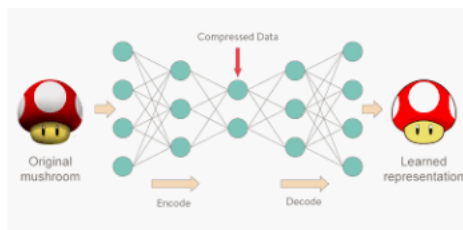
다변량

비지도 학습

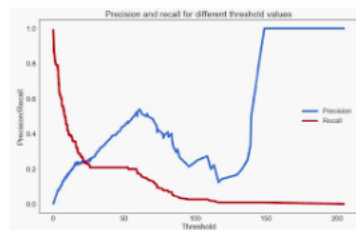
Autoencoder



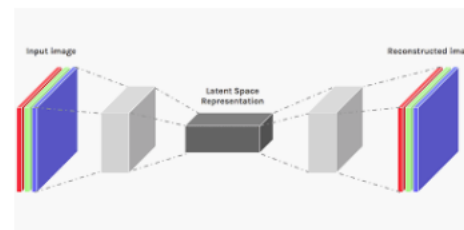
Deep Autoencoders using Tensorflow ...
towardsdatascience.com



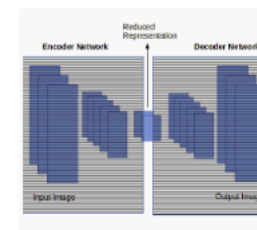
Deep Autoencoders using Tensorflow ...
towardsdatascience.com



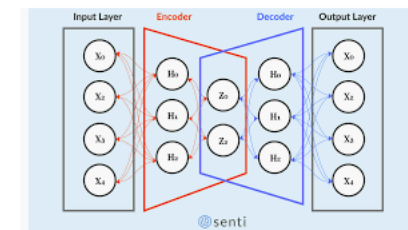
Fraud Detection Using Autoencoders in ...
datascience.com



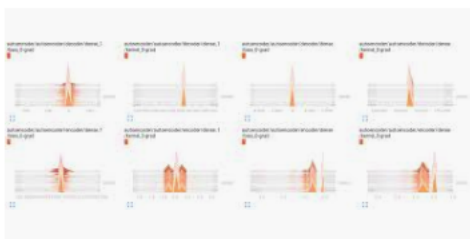
Autoencoders — Introduction and ...
towardsdatascience.com



color images in Tensorflow fra...
mc.ai



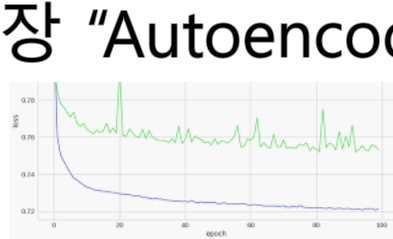
TensorFlow 2.0 – Abien Fred Agarap ...
medium.com



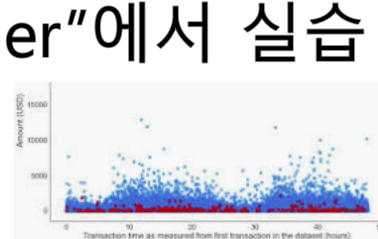
Autoencoder loss is not decreasing (and ...
stackoverflow.com

	Label	reconstruction_e
count	910.000000	910.000000
mean	0.026374	0.009409
std	0.160332	0.019738
min	0.000000	0.000156
25%	0.000000	0.001822
50%	0.000000	0.004485
75%	0.000000	0.009801
max	1.000000	0.154782

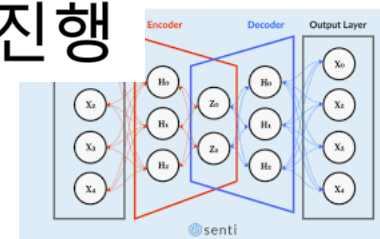
Neural Networks for Anomaly (...
blog.goodaudience.com



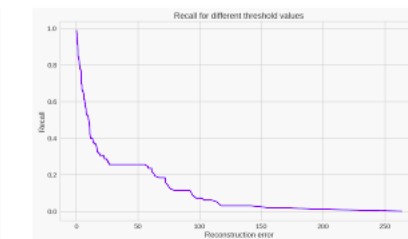
Credit Card Fraud Detection using ...
medium.com



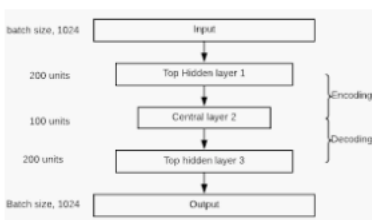
Fraud Detection Using Autoencoders in ...
datascience.com



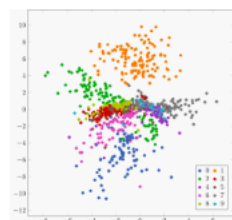
Autoencoder in TensorFlow ...
senti.ai



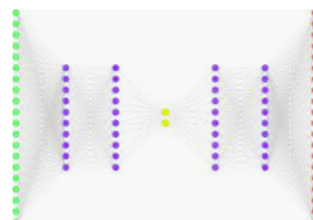
Credit Card Fraud Detection using ...
medium.com



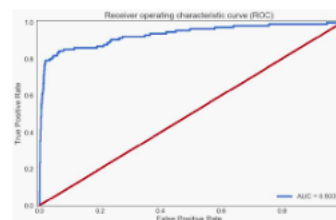
TensorFlow Autoencoder: Deep Learning ...
guru99.com



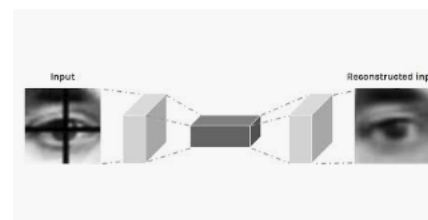
Simple MNIST Autoencode...
gertjanvandenburgh.com



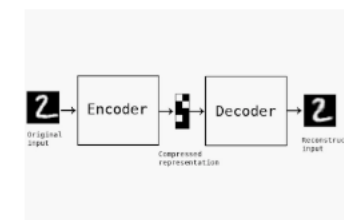
Simple MNIST Autoencoder in Tensor...
gertjanvandenburgh.com



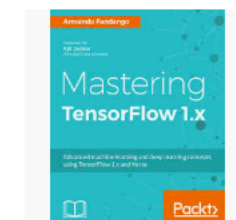
Fraud Detection Using Autoencoders in ...
datascience.com



Autoencoders — Introduction and ...
towardsdatascience.com



Building Autoencoders in Keras
blog.keras.io



Denosing autoencoder in ...
subscription.packtpub.com

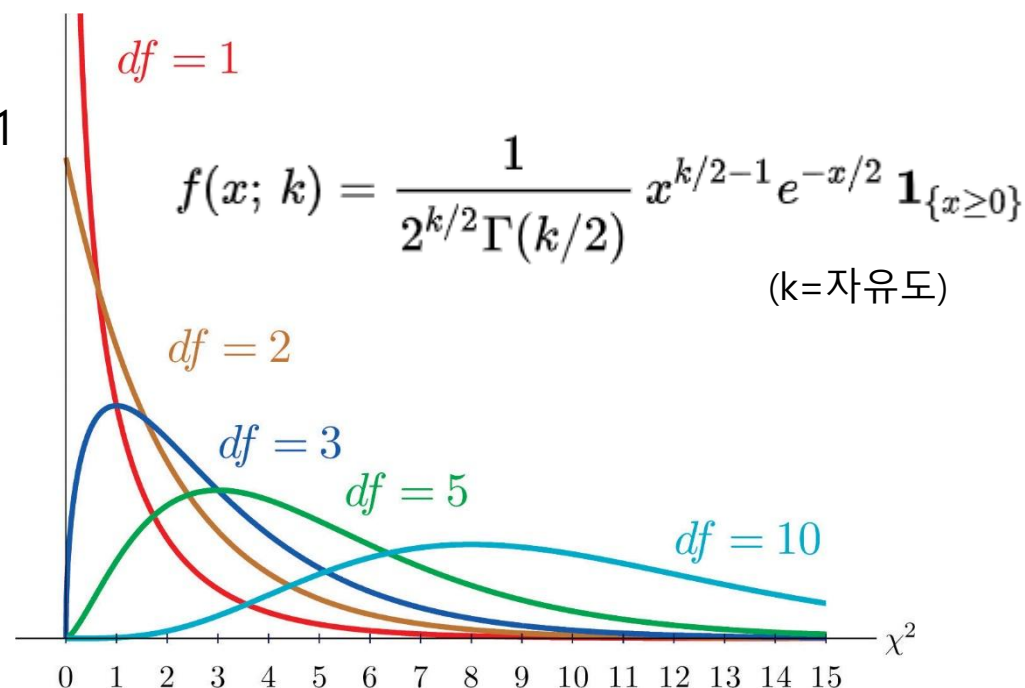
4장 "Autoencoder"에서 실습 진행

유의차 분석

Category by Category

Cross-Tabulation

- Cross-Tabulation = Contingency Table
- Chi-square test : 변수의 독립성 검정(independence of variables)
 - chi-square statistic : 기대값과 관찰값의 차이가 클수록 Chi-square 통계량의 값이 커짐
 - p-value : chi-square statistic의 유의 확률
 - degree of freedom : 자유도(그래프의 df) , 범주의 수 - 1
- Category by Category
 - 설비별 Good/NG 빈도 차이
 - 전고정A 설비와 자공정B 설비의 Lot 진행수 차이
- Analytics
 - One-Way Chi-square Test
 - N-Way Chi-square Test



유의차 분석

Category by Category

Cross-Tabulation

- `scipy.stats.chi2_contingency`
 - parameters
 - `observedarray_like`
 - The contingency table. The table contains the observed frequencies (i.e. number of occurrences) in each category. In the two-dimensional case, the table is often described as an "R x C table".
 - `correctionbool`, optional
 - If True, and the degrees of freedom is 1, apply Yates' correction for continuity. The effect of the correction is to adjust each observed value by 0.5 towards the corresponding expected value.
 - `lambda_float` or `str`, optional.
 - By default, the statistic computed in this test is Pearson's chi-squared statistic [2]. `lambda_` allows a statistic from the Cressie-Read power divergence family [3] to be used instead. See `power_divergence` for details.
 - Returns
 - `chi2float` : The test statistic.
 - `pfloat` : The p-value of the test
 - `dofint` : degrees of freedom
 - `expectedndarray`, same shape as `observed` : The expected frequencies, based on the marginal sums of the table.

유의차 분석

Category by Category

Cross-Tabulation

Chi-Square를 이용한 품질 유의차 분석

- 서로 다른 Line 또는 설비를 통과한 Lot내 Good/NG 빈도 산출
- Cross-Table(=Contengency Table) Formatting
- Chi-square Test를 이용한 설비 유의차 분석

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

Case A

기대확률			
구분	설비 A	설비 B	계
OK	0.3	0.3	0.6
NG	0.2	0.2	0.4
계	0.5	0.5	1.0

$\chi^2 =$

검사내역			
구분	설비 A	설비 B	계
OK	30	30	60
NG	20	20	40
계	50	50	100

Case B

검사내역			
구분	설비 A	설비 B	계
OK	25	35	60
NG	25	15	40
계	50	50	100

$\chi^2 =$

Case C

검사내역			
구분	설비 A	설비 B	계
OK	10	50	60
NG	0	40	40
계	50	100	150

$\chi^2 =$

기대빈도

구분	설비 A	설비 B	계
OK	30	30	60
NG	20	20	40
계	50	50	100

실질빈도 - 기대빈도

구분	설비 A	설비 B
OK	0	0
NG	0	0

Cell-Chisquare

	설비 A	설비 B
OK	0	0
NG	0	0

실질빈도 - 기대빈도

구분	설비 A	설비 B
OK	-5	5
NG	5	-5

Cell-Chisquare

	설비 A	설비 B
OK	0.8	0.8
NG	1.25	1.25

실질빈도 - 기대빈도

구분	설비 A	설비 B
OK	-20	20
NG	20	-20

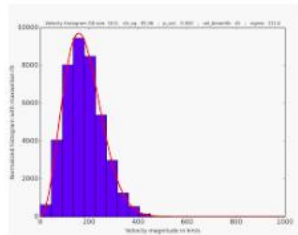
Cell-Chisquare

	설비 A	설비 B
OK	13.3	13.3
NG	20	20

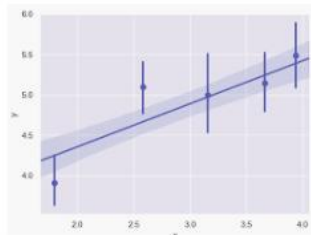
유의차 분석

Category by Category

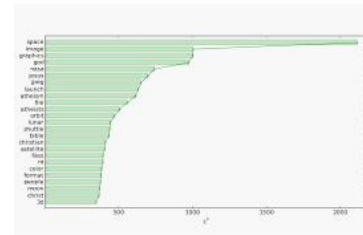
Cross-Tabulation



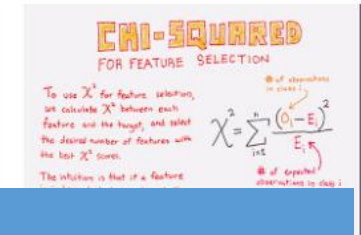
chi squared test on a normalized ...
stats.stackexchange.com



Python Linear Regression | Chi-Squa...
data-flair.training



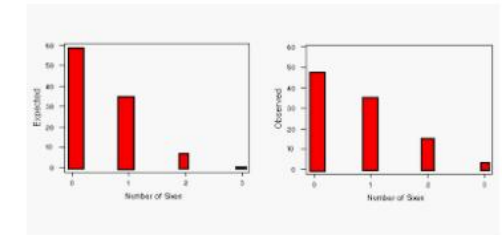
Terms selection with chi-square
glowingpython.blogspot.com



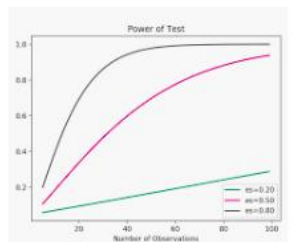
실습

	First year dropout	Second year dropout	Third year dropout	Fourth year dropout	Graduated
Observed	543	145	210	110	992
Expected	$6.2 \times 2000 = 400$	$0.1 \times 2000 = 200$	$0.15 \times 2000 = 300$	$0.05 \times 2000 = 100$	$0.5 \times 2000 = 1000$

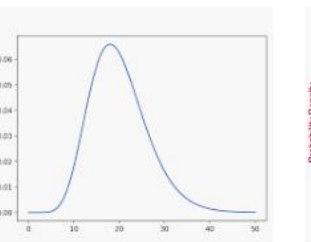
Statistics: Pearson's chi-squared test ...
medium.com



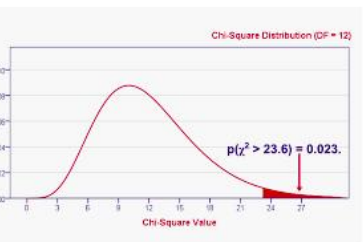
Chi-Square Goodness of Fit Test
stat.yale.edu



Statistical Power and Power Analy...
machinelearningmastery.com



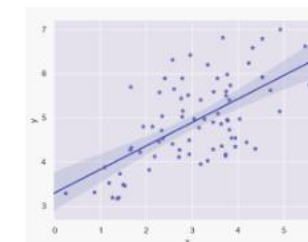
Statistical Data Distributions
machinelearningmastery.com



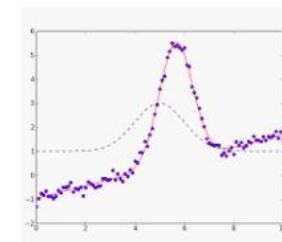
Chi-Square Test - Beginners Tutorial ...
spss-tutorials.com

id	name	age	sex	height	weight	hair	eyes	skin	hair_color	eye_color	skin_color	hair_type	eye_type	skin_type
1	Michael, Mr. Owen Harris	42.0	1	1.73	70.0	1	1	1	1	1	1	1	1	1
2	Curtis, Mrs. John Graham (Graham)	38.0	1	1.65	65.0	1	1	1	1	1	1	1	1	1
3	Indira, Mrs. Lata	35.0	1	1.60	60.0	1	1	1	1	1	1	1	1	1
4	Patricia, Mrs. Josephine (Pat) Mary	30.0	1	1.55	55.0	1	1	1	1	1	1	1	1	1
5	John, Mr. William Henry	25.0	1	1.70	70.0	1	1	1	1	1	1	1	1	1

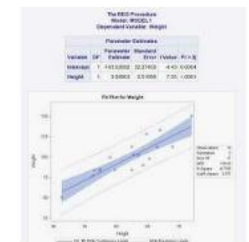
Chi Square Feature Selection in Python ...
insightsbot.com



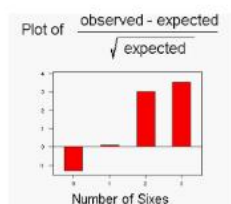
Python Linear Regression | Chi-Squa...
data-flair.training



Modeling Data and Curve Fitting ...
cars9.uchicago.edu



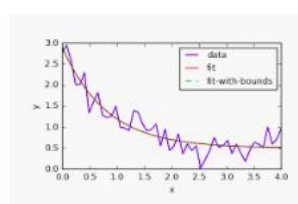
prediction limits in scatter ...
stackoverflow.com



Chi-Square Goodness of F...
stat.yale.edu



Chi-Square Test - Beginners Tutorial ...
spss-tutorials.com



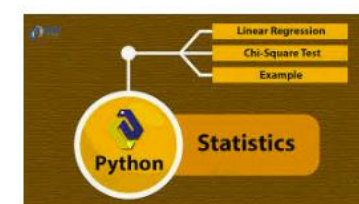
scipy.optimize.curve_fit — SciPy v0...
docs.scipy.org



Chi-Square Tests with Die Roll Data ...
towardsdatascience.com



Python Linear Regression | Chi-...
data-flair.training



Python Linear Regression | Chi-Square ...
data-flair.training

Chi-Square Test: red, yellow, green, blue					
Expected counts are printed below observed counts					
1	red	yellow	green	blue	Total
1	50.00	10.00	20.00	20.00	100.00
2	150.00	30.00	60.00	60.00	300.00
Total	200.00	40.00	80.00	80.00	400.00
Chi-Sq = 18.909 + 1.600 + 5.900 + 28.600 + 6.900 + 0.133 + 1.667 + 9.600 = 71.200					
df = 3, P-Value = 0.000					

Data Analysis - Chi-squared test for ...
learntech.uwe.ac.uk

04. 품질 최적화

유의차 분석

Parameter by Category

ANOVA

Parameter 간 유의차 분석

- 서로 다른 Line, 설비 또는 설비Parameter를 통과한 Lot 추적
- Line/설비를 통과한 Lot의 Defect Type별 Good/Bad 판정 및 불량율 산출
- 불량율 유의차 판별

Lot 공정 투입

lot 1-A



lot 2-A



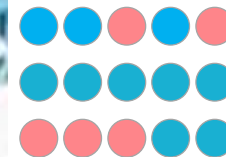
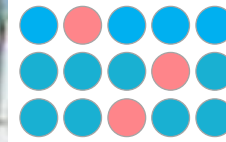
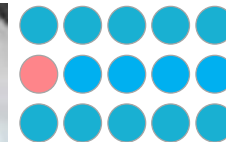
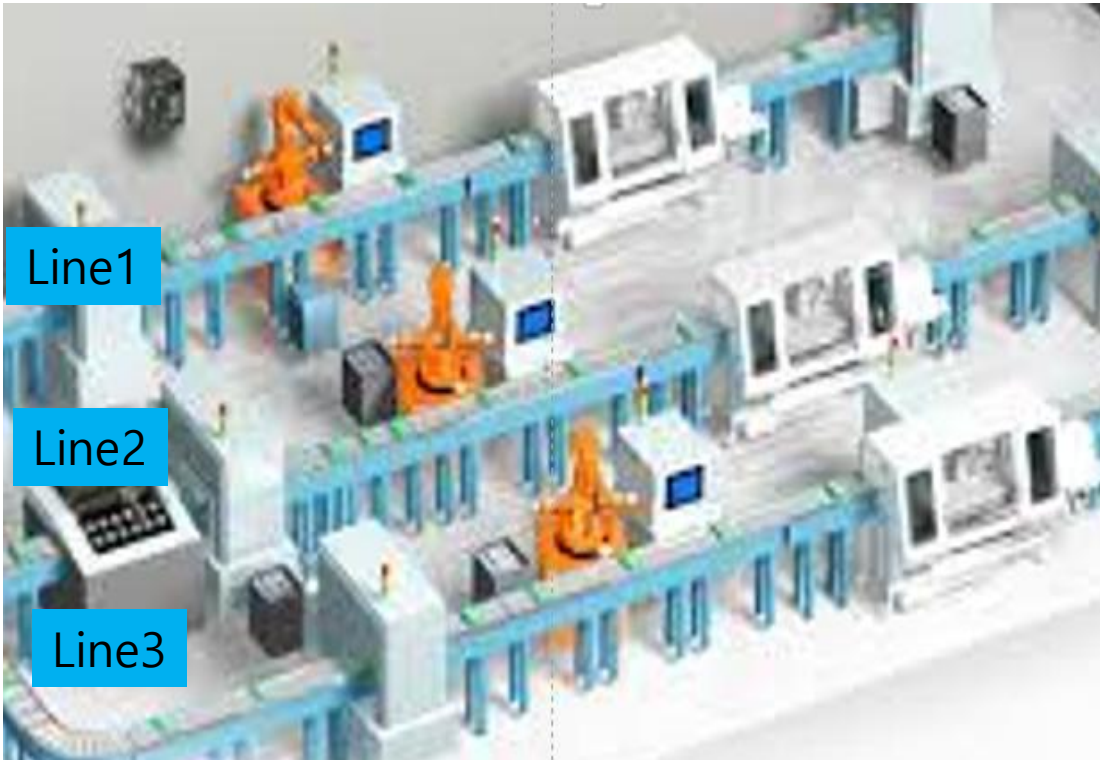
lot 3-A



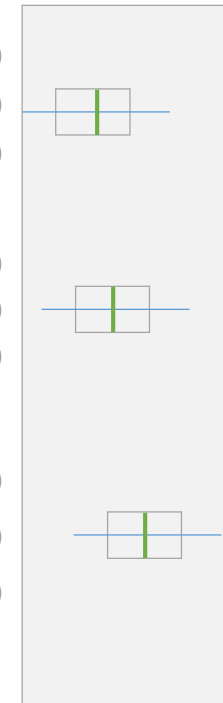
Line1

Line2

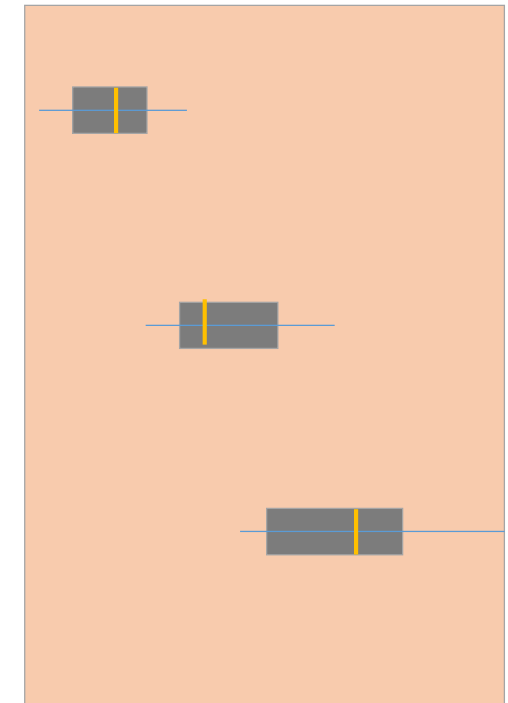
Line3



case A



case B



유의차 분석

Parameter by Category

ANOVA

집단간 측정값의 분산을 고려한 평균의 차이 분석

- 집단 : 공정, 장비, 레시피, 작업자등
- 측정 : 온도, 압력, 두께 등
- 예
 - 동일 역할 장비간의 재공 상태 분석

- `statsmodels.stats.anova.anova_lm`
 - Anova table for one or more fitted linear models.
- Parameters
 - `argsfitted` linear model results instance
 - One or more fitted linear models
 - `scalefloat`
 - Estimate of variance, If None, will be estimated from the largest model. Default is None.
 - `teststr` {"F", "Chisq", "Cp"} or None
 - Test statistics to provide. Default is "F".
 - `typstr` or int {"I","II","III"} or {1,2,3}
 - The type of Anova test to perform. See notes.
 - `robust`{None, "hc0", "hc1", "hc2", "hc3"}
 - Use heteroscedasticity-corrected coefficient covariance matrix. If robust covariance is desired, it is recommended to use hc3.

유의차 분석

Parameter by Category

ANOVA

집단간 측정값의 분산을 고려한
평균의 차이 분석

- 집단 : 공정, 장비, 레시피, 작업자 등
- 측정 : 온도, 압력, 두께 등
- 예
 - 동일 역할 장비간의 재공 상태 분석

- Returns
 - anovaDataFrame
 - When args is a single model, return is DataFrame with column
 - sum_sqfloat64 : Sum of squares for model terms.
 - dffloat64 : Degrees of freedom for model terms.
 - Ffloat64 : F statistic value for significance of adding model terms.
 - PR(>F)float64 : P-value for significance of adding model terms.
 - When args is multiple models, return is DataFrame with columns:
 - df_residfloat64 : Degrees of freedom of residuals in models.
 - ssrfloat64 : Sum of squares of residuals in models.
 - df_difffloat64 : Degrees of freedom difference from previous model in args
 - ss_difffloat64 : Difference in ssr from previous model in args
 - Ffloat64 : F statistic comparing to previous model in args
 - PR(>F): float64 : P-value for significance comparing to previous model in args

유의차 분석

Parameter by Category

ANOVA

ANOVA Mode(Case 1) - Treatment에 따른 Parameter값의 차이

	df	sum_sq	mean_sq	F	PR(>F)
C(treatment)	2.0	15515.766414	7757.883207	3.711336	0.043589
Residual	19.0	39716.097222	2090.320906	NaN	NaN

ANOVA Model(Case 2) – Treatment, obser, fetu*observer(교호작용)에 따른 Parameter값의 차이

	df	sum_sq	mean_sq	F	PR(>F)
C(fetus)	2.0	324.008889	162.004444	2113.101449	1.051039e-27
C(observer)	3.0	1.198611	0.399537	5.211353	6.497055e-03
C(fetus):C(observer)	6.0	0.562222	0.093704	1.222222	3.295509e-01
Residual	24.0	1.840000	0.076667	NaN	NaN

04. 품질 최적화

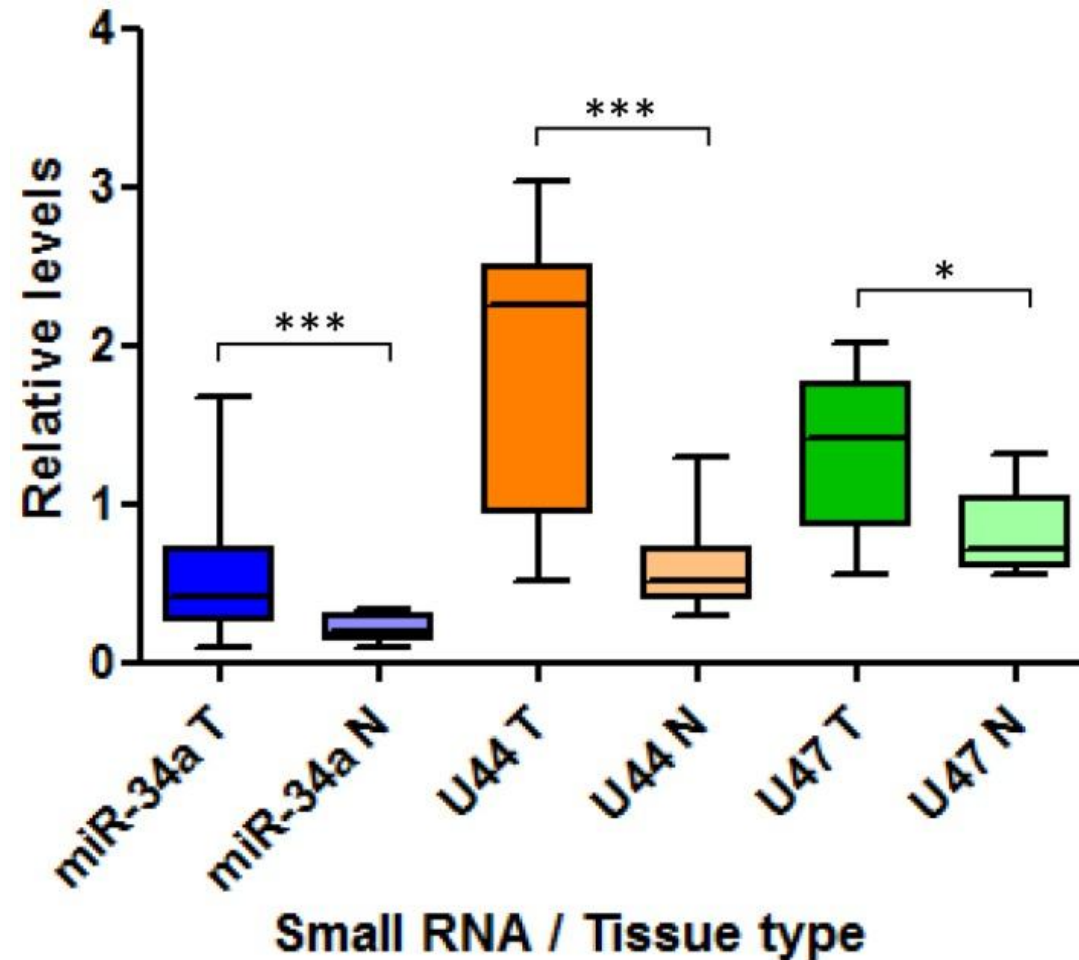
유의차 분석

Parameter by Category

ANOVA

시각화 및 Parameter의 유의차

표현	신뢰수준	유의수준
***	99%	$p < 0.01$
**	95%	$p < 0.05$
*	90%	$p < 0.1$



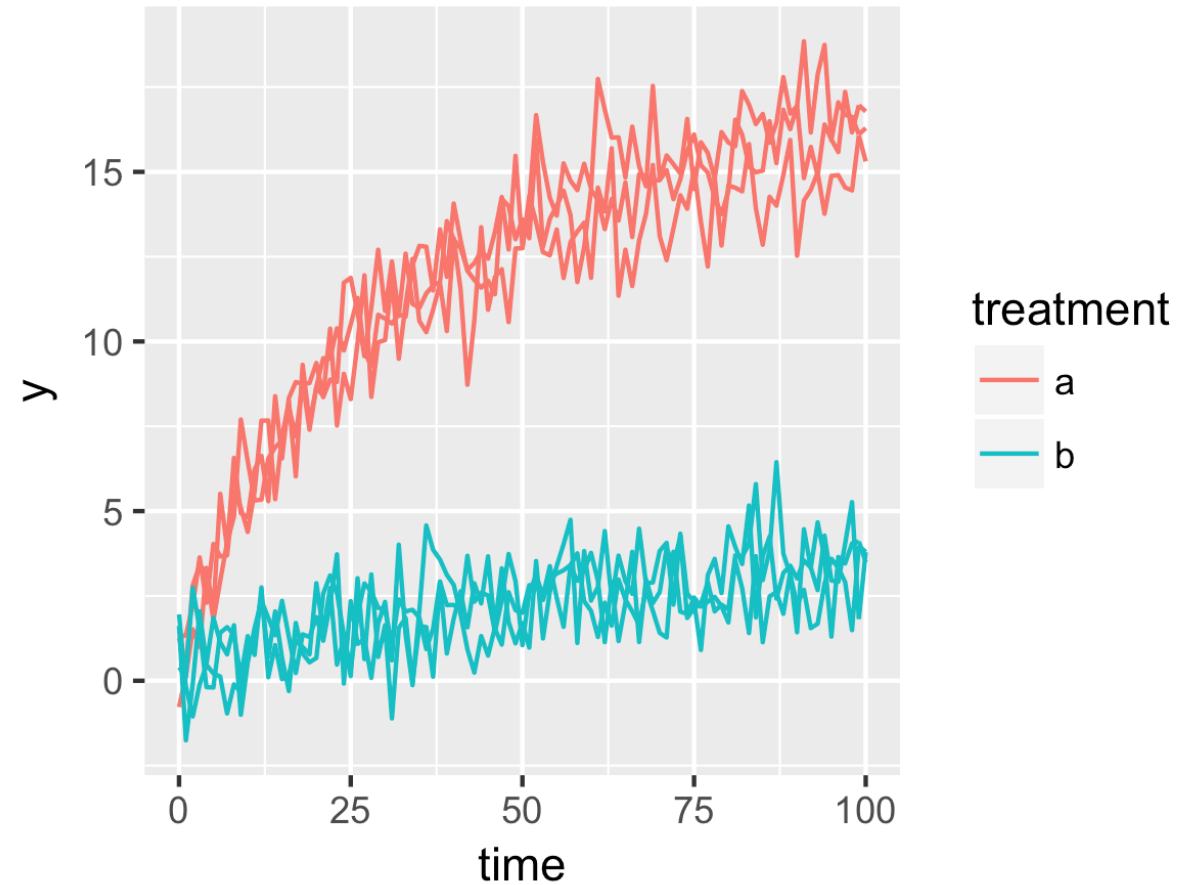
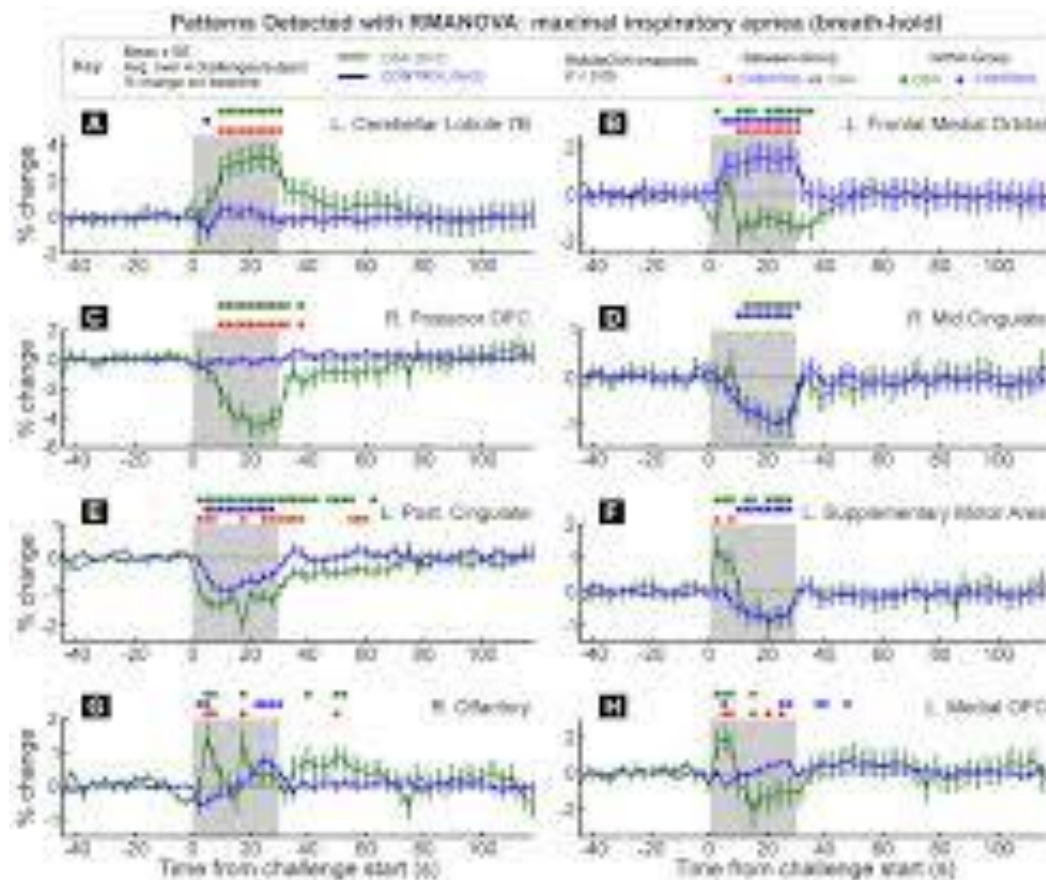
04. 품질 최적화

유의차 분석

Parameter by Category

ANOVA

■ ANOVA의 응용 : 시계열 구간별 유의차 검토

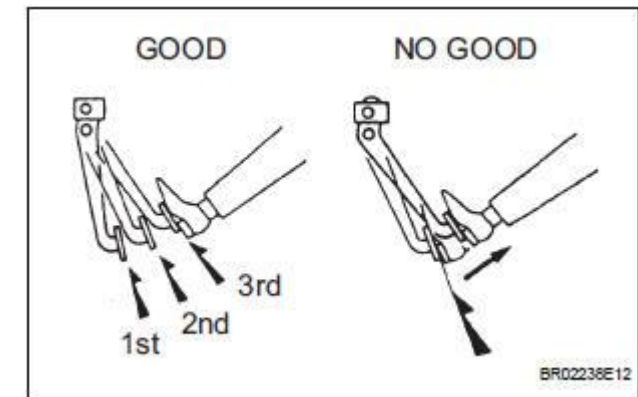
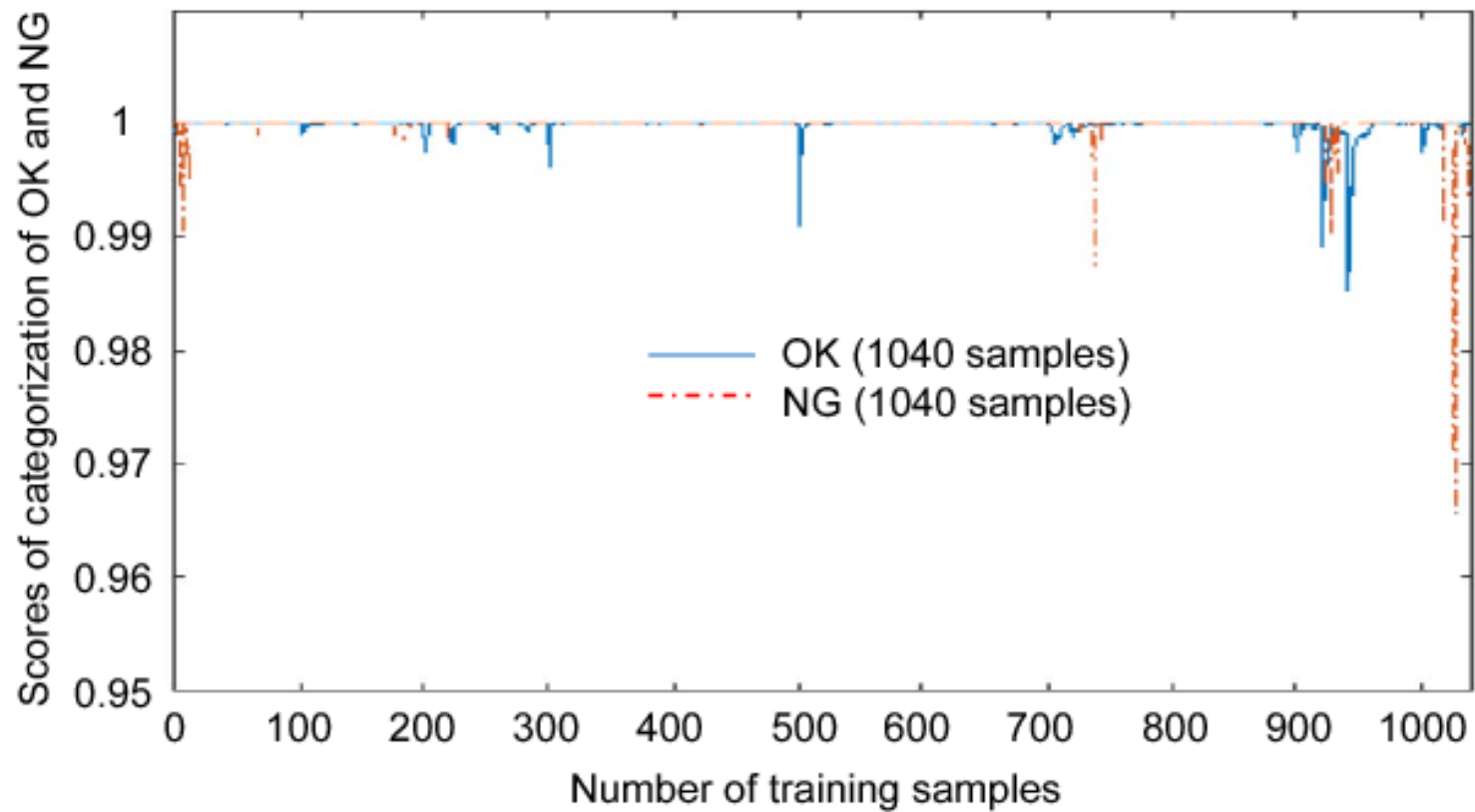


유의차 분석

Parameter by Category

ANOVA

■ ANOVA의 응용 : Good/NG의 유의차 분석

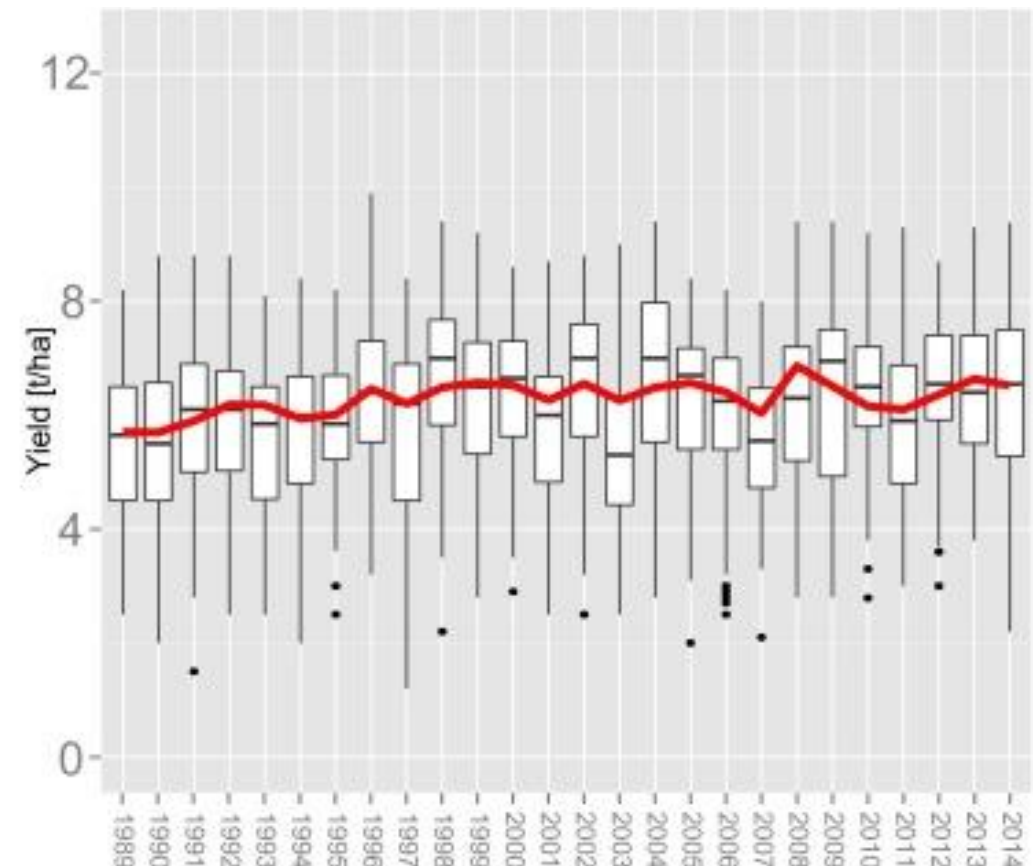
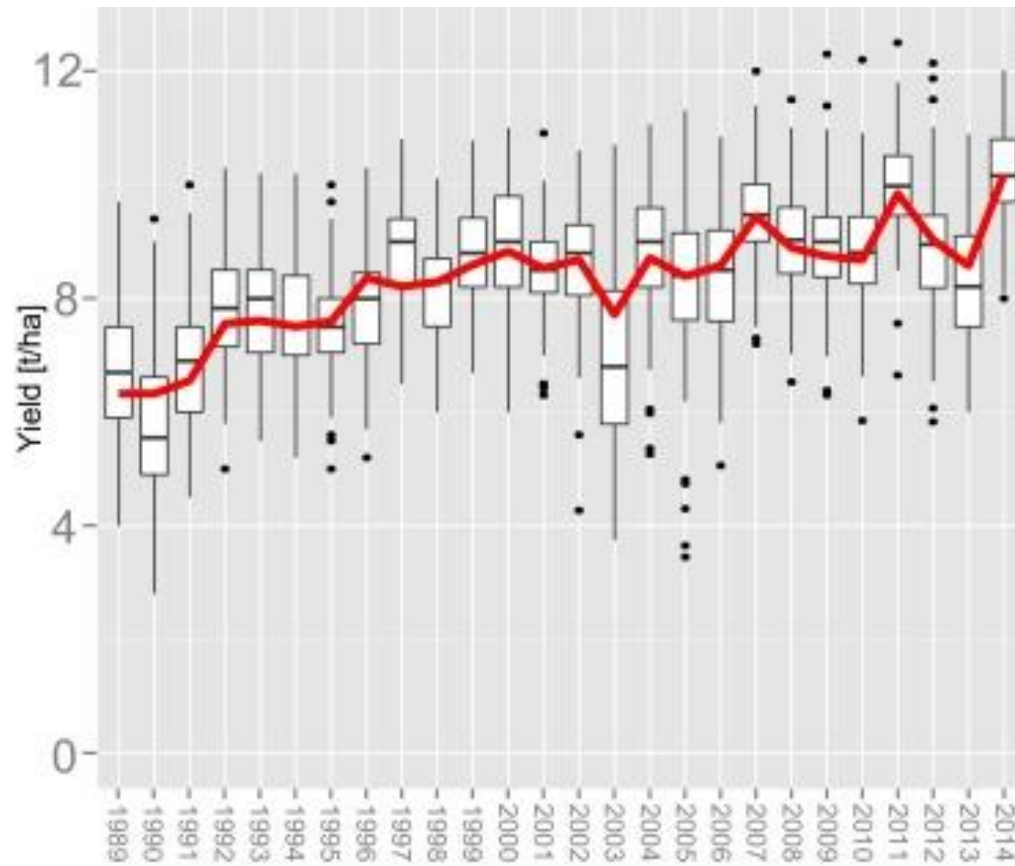


유의차 분석

Parameter by Category

ANOVA

■ ANOVA의 응용 : 모델 결과와 병행하여 Boxplot의 이상치 검토

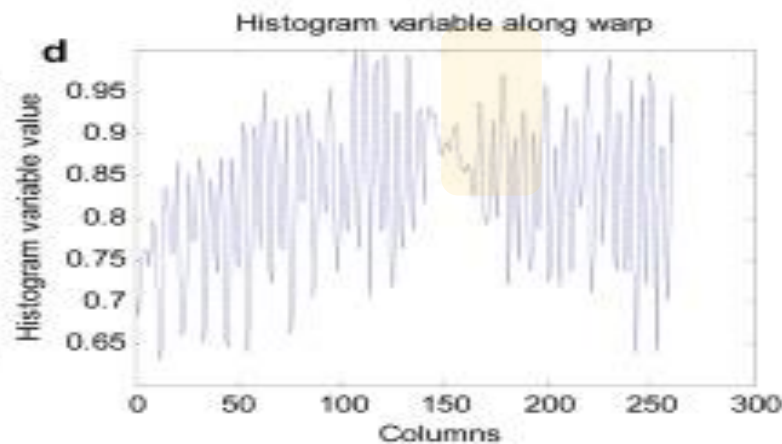
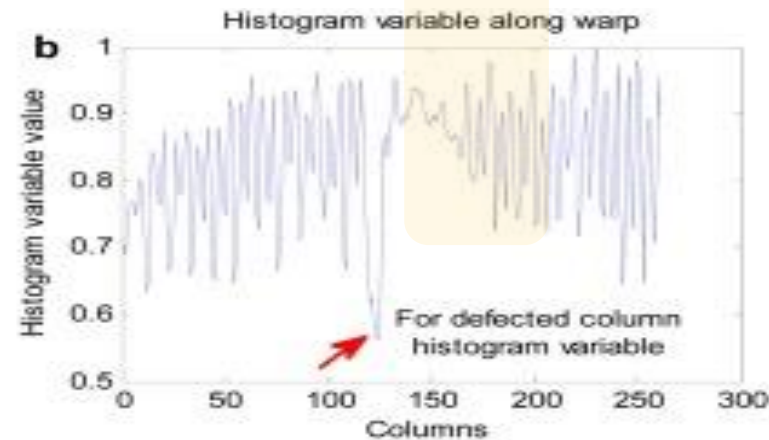
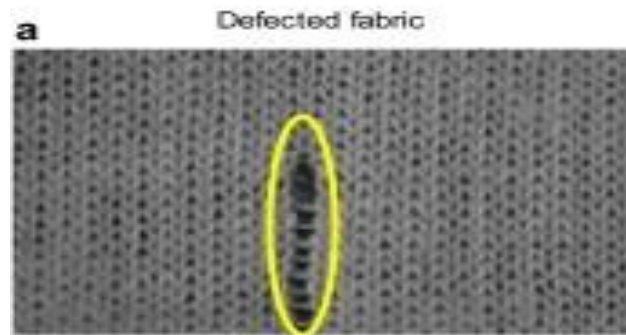


유의차 분석

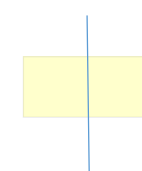
Parameter by Category

ANOVA

■ ANOVA의 응용 : Good/NG의 이상치 식별



■ outlier

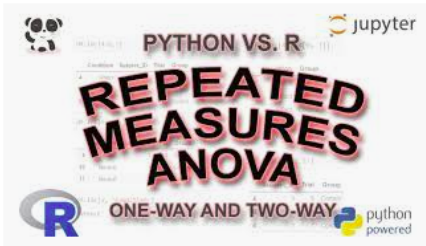


04. 품질 최적화

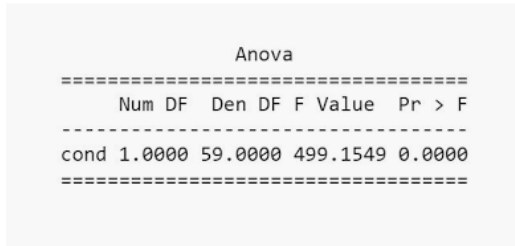
유의차 분석

Parameter by Category

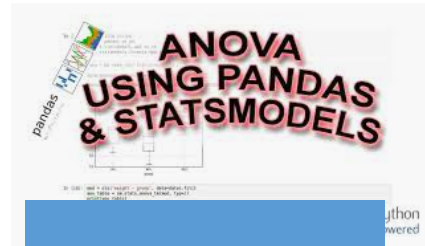
ANOVA



Repeated measures ANOVA using Python ...
youtube.com



Repeated Measures ANOVA in Python using ...
marsja.se



2-way ANOVA in Python statsmodels ...
stats.stackexchange.com

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9.785 ^a	5	1.957	26.021	.000
Intercept	161.687	1	161.687	2149.844	.000
Factor_1	5.943	1	5.943	79.026	.000
Factor_2	3.811	2	1.906	25.339	.000
Factor_1 * Factor_2	.039	2	.020	.262	.770
Error	3.760	50	.075		
Total	175.803	56			
Corrected Total	13.545	55			

a. R Squared = .722 (Adjusted R Squared = .695)

2-way ANOVA in Python statsmodels ...
stats.stackexchange.com

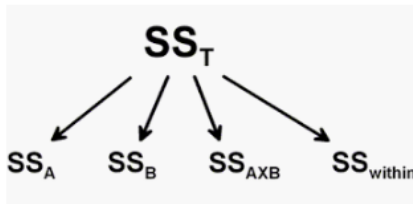
Tests of Between-Subjects Effects					
Dependent Variable: Effect					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	319.276 ^a	3	106.425	117.692	.000
Intercept	72.981	1	72.981	80.706	.000
Fiction	67.557	1	67.557	74.709	.000
Condition	11.514	1	11.514	12.733	.000
Fiction * Condition	10.776	1	10.776	11.917	.001
Error	520.862	576	.904		
Total	1116.000	580			
Corrected Total	840.138	579			

a. R Squared = .380 (Adjusted R Squared = .377)

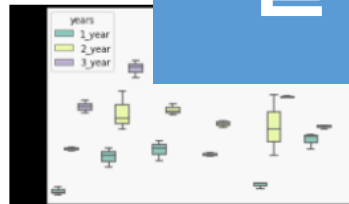
ANOVA results to SPSS ...
stackoverflow.com

ANOVA table	SS	DF	MS	F (DF1, DF2)	P-value
Interaction	0.03944	2	0.01972	F (2, 50) = 0.2622	P = 0.7704
Factor 2	3.811	2	1.906	F (2, 50) = 25.34	P < 0.0001
Factor 1	0.043	1	0.043	F (1, 50) = 79.03	P < 0.0001
Residual	3.760	50	0.07521		

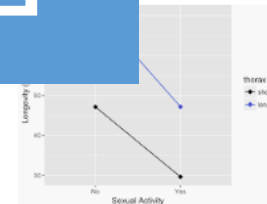
2-way ANOVA in Python statsmodels ...
stats.stackexchange.com



two-way ANOVA with Python ...
marsja.se



ANOVA using Python | Renesh Bedre
reneshbedre.github.io



Statistical Models in R: Day 1
asda.stat.ubc.ca

```
> anova(lm(StressReduction ~ Treatment * Age, data=twoWayComparisons))
Analysis of Variance Table

Response: StressReduction
Df Sum Sq Mean Sq F value Pr(>F)
Treatment 2 18 9 0.001953 ***
Age 2 362 81 1e+05 ***
Treatment:Age 4 0 0 2.250e-30 1.000000
Residuals 18 18 1
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

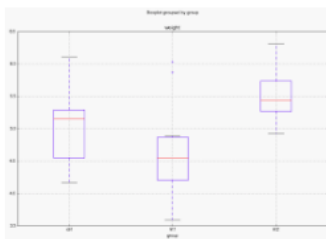
How to get the degree of freedom in anova
jenniferdanos.com

Multiple Comparison of Means - Tukey HSD, FWER=0.05						
group1	group2	meandiff	lower	upper	reject	
high	low	-1.8	-4.1651	0.5651	False	
high	placebo	-2.8	-5.1651	-0.4349	True	
low	placebo	-1.0	-3.3651	1.3651	False	

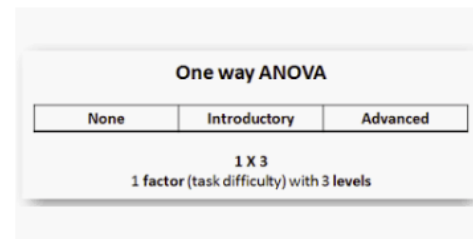
Analysis of Variance (ANOVA) | Python ...
pythonfordatascience.org

len	supp	dose
4.2	VC	0.5
11.5	VC	0.5
7.3	VC	0.5
5.8	VC	0.5
6.4	VC	0.5
10.0	VC	0.5
11.2	VC	0.5
11.2	VC	0.5
5.2	VC	0.5
700 x 380		

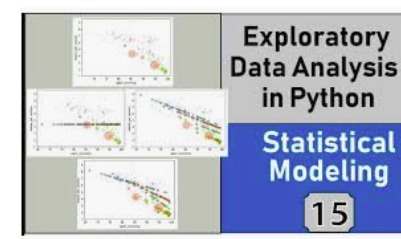
Anova in Python | plotly
plotly



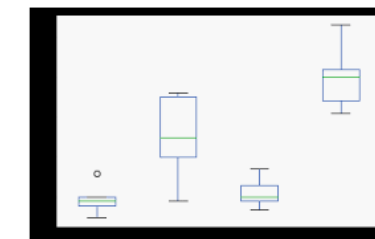
one-way ANOVA with Python ...
marsja.se



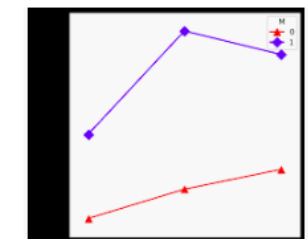
anova-one
pytolearn.csd.auth.gr



Anova, Fitting Models To Data ...
youtube.com



ANOVA using Python | Renesh Bedre
reneshbedre.github.io



interactions_anova
statsmodels.org

THANK YOU

감사합니다