

2.2 잡음(noisy data) 제거 - 실무예제



다음은 2016년 항만별 선박 입출항 현황이다. 선박 수 속성은 구간화(Binning) 기법으로 평활화하고, 선박 톤수 속성은 선박 수 속성과의 회귀(regression) 분석을 통하여 평활화시키시오.

◎ 데이터 파일 : ch2-2(선박입출항).csv

◎ 원본 튜플 수 : 30개

항만	입항선박수	입항선박톤수	출항선박수	출항선박톤수
부산	7,301	105,138,280	7,409	103,857,903
인천	2,715	30,716,710	2,716	30,779,186
평택.당진	1,558	23,153,226	1,536	22,778,109
경인항	28	126,236	27	128,344
동해.목호	552	4,202,603	546	4,039,929

출처 공공데이터포털(www.data.go.kr)

☞ 해답은 [ch2-2.ipynb](#) 참고