**Question 2 분석결과**

**텍스트, 스크린샷, 폰트, 그래프이(가) 표시된 사진

자동 생성된 설명**

위 그래프를 보면, 시행 횟수별로 (각 그래프의 title이 시행 횟수를 나타낸다) 얼만큼의 빈도가 측정 되었는지를 나타낸 히스토그램이다(코드 상에서는 bar(막대그래프) 형식으로 구현했지만, 결과에 따라서 빈도 수를 Dictionary 형태로 저장 후, 그 결과값 자체를 막대 그래프 형식으로 저장했기에 히스토그램과 같은 역할을 한다). 참고적으로 시행 결과는 시행 횟수별로 random하게 1~6까지의 숫자를 만들어 나온 결과이다. 따라서, .py 실행 시마다 Q2 폴더 내부에 있는 dice\_result.csv의 값은 달라질 것이며, 그에 따른 위의 그래프의 분석 값 또한 달라질 수 있다는 것을 참고하자. 이번 분석에서는 위 히스토그램을 기준으로 분석할 예정이다.

히스토그램은 영역별로의 데이터 특성과 전체적인 데이터 분포를 시각적으로 파악할 때 유리한 그래프이다. 위 히스토그램을 살펴보면 시행 횟수가 늘어날수록 주사위에서 1~6이 각각 나오는 횟수의 시행 횟수 대비 비율이 비슷해 진다는 것을 시각적으로 쉽게 확인할 수 있다. 히스토그램의 막대 길이가 시행 횟수가 늘어날수록 서로 비슷해 지는 것을 확인할 수 있기 때문이다.

이를 좀 더 정확히, 수치적으로 분석해 봐도 똑같은 결론을 지을 수 있다. 참고를 위해 각 히스토그램마다 1~6이 나온 빈도수를 그래프에 표기했다.

시행 횟수별로 1~6까지 나오는 빈도의 비율을 정리해보면 아래 표와 같이 정리해볼 수 있다.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 나온 숫자 | 100회 시행 | 1,000회 시행 | 10,000회 시행 | 100,000회 시행 |
| 1 | 0.19 | 0.18 | 0.19 | 0.19 |
| 2 | 0.16 | 0.19 | 0.19 | 0.18 |
| 3 | 0.15 | 0.17 | 0.19 | 0.19 |
| 4 | 0.15 | 0.17 | 0.18 | 0.19 |
| 5 | 0.13 | 0.18 | 0.18 | 0.18 |
| 6 | 0.22 | 0.21 | 0.18 | 0.19 |

*(위 수치는 편의를 위해 소수 셋째자리에서 반올림한 값들임)*

수치적으로 확인한 위 표를 토대로, 시행 횟수가 늘어날수록 0.2의 값에 모두 수렴하며 전체 시행 횟수 대비 각각의 숫자가 나오는 비율이 비슷해 지는 것을 확인할 수 있다.