

인스타그램의 게시물과 해시태그를 이용한 은평한옥마을의 특징 분석

연구자: 이준석, 권희선, 김명서
손준영, 신우현, 이호현
한도윤, 황두현

《 초 록 》

본 연구는 SNS 인스타그램에서 하나의 해시태그를 검색하여 그 해시태그를 포함하고 있는 모든 게시글의 다른 해시태그, 게시 시점, 좋아요 수, 조회 수, 댓글 수 데이터를 Instaloader를 활용하여 수집하였다. 다음의 방식으로 수집한 데이터들은 .json파일로 저장하여 이에 대한 파싱 작업을 거쳐 빈도 분석을 진행하였다. 또한, 다음의 방식으로 수집한 데이터들을 도표와 워드 클라우드를 활용하여 시각화를 진행하였다. 해시태그 '은평한옥마을'의 경우, 2014년 11월 5일 게시물부터 2022년 7월 12일 게시물까지 총 47,227개의 게시물에 대한 분석을 진행하였고 해시태그 '은평한옥마을카페'의 경우에는 2016년 12월 17일 게시물부터 2022년 10월 19일까지 작성된 2915개의 게시물에 대해 분석을 진행하였다.

본 연구는 인스타그램을 통하여 은평한옥마을에 대한 특징을 분석하였고 이를 통해 은평한옥마을에서는 북촌한옥마을에 비해 전체적인 활성화도가 낮음을 알 수 있었다. 또한 은평한옥마을의 해시태그를 카페, 음식점, 풍경, 방문 동기, 그리고 은평한옥마을과 관련이 없는 태그 즉, 5가지로 나누어 은평한옥마을의 특징을 분석하였다.

목 차

I. 서론	03
1. 연구 동기	03
2. 연구 목적	04
II. 이론적 배경	04
1. 인스타그램 특성	04
2. 웹 스크래핑 및 시각화	06
III. 연구 방법	06
1. 연구내용 및 방법	06
IV. 결과 및 고찰	07
1. 연구 결과	07
2. 결과 해석	11
V. 결론 및 제언	13
VI. 참고 문헌	13
VII. 부록	14
1. Instaloader를 활용하여 웹 스크래핑	14
2. .json 파일에 대한 파싱을 위한 Python 파일	15

I. 서론

1. 연구 동기

인터넷의 발전에 따라, 우리 사회는 점차 상호 연결적인 면모를 띄고 있다. 이에, 방문 및 관광·레저의 형태와 그 특징이 달라지고 있다. 과거에는 정보의 수집과 교환이 어려움에 따라 정해진 형태의 지역 동선을 중심으로 움직였다. 특히, 이미 해당 지역에 대한 지식을 가지고 있는 전문가를 대동해 이동 및 방문하는 경우가 많았으며, 이러한 특징들의 결과로 전반적으로 지역 주민과 외부 방문객의 방문 장소와 경험이 큰 괴리를 띄었으며, 전통적인 방문 동선에서 제외된 지역은 외부 방문객을 맞는 경우가 거의 없었다.

그러나, 최근 몇 년 동안 기술 개발에 힘입어 정보의 개방성이 증대되면서, 단체 방문객이 아닌 개인 방문객들이 스스로 해당 지역에 대한 정보를 수집하고, 이를 바탕으로 본인만의 새로운 이동 동선을 구성하는 경우가 늘어나고 있다. 특히, 기관이나 기업에 의한 일 대 다의 연결이 아니라, 개인과 개인 사이의 일대일 연결의 성장이 괄목할 만하다.

이러한 개인과 개인의 연결은 소셜 네트워크 서비스(SNS)로 대표된다. 스마트폰의 보급과 잇따른 SNS 이용자의 증가로, 사람들 간의 연결 고리는 매체상에서 더욱 활발해지고 있다. 자신과 이질적인 외부 사용자와 접근할 수 있는 환경인 만큼, SNS의 주 이용 목적은 커뮤니케이션과 콘텐츠의 생산 및 공유라고 할 수 있다. 결과적으로, 사람들은 SNS에 자신들의 생각을 공유하여 자신의 영향력을 증대할뿐더러, 타인과의 연결에서 결부되는 콘텐츠의 영향 아래에 있게 된다. 이러한 경향은 미래 사회를 이끌 젊은 세대에서 그 특징이 두드러지며, 소위 ‘디지털 네이티브’라는 수식어로 불리는 현재의 20·30세대는 정보를 그 어떠한 대중 매체보다 SNS상에서 더 많이 얻는다. 이러한 SNS 산업의 증가세에 따라 정보량이 증가하면서 최근 유수의 기관과 기업에서 ‘신세대’ 맞춤형 홍보전략을 마련하기 위해 SNS상의 정보를 분류하는 해시태그에 주목하고 있으며, 이를 이용해 대상의 특징을 파악하는 연구가 여러 기관에서 진행되고 있다. 따라서, SNS의 게시물과 해시태그를 분석함으로써 방문지의 특징을 분석할 수 있으며 이를 이용하여 부속적인 연구를 진행하는데 일조할 수 있는 것이다.

이렇게 인터넷의 발달에 따른 개인의 독립성으로 말미암아, 기존에 방문객의 수가 서울의 다른 지역에 비해 적던 은평구를 찾는 사람의 수도 증가하고 있다. 이 중심에는 은평구의 특징이자 서울시 내의 유일한 계획형 주거 한옥 단지로 방문객들에게는 특별한 경험을 제공하는 은평한옥마을이 있다. 이에, 본 연구는 은평한옥마을에 대한 온라인상의 인지도 변화를 활용하여 은평 한옥마을의 특성을 분석할 수 있다고 생각하였으며, 이에 대한 도구로 타 기관에서도 자주 사용되는 인스타그램 해시태그 분석을 선정하였다.

2. 연구 목적

본 연구는 상술하였듯이 늘어나는 인터넷 영향력을 활용하여 은평한옥마을의 특징을 분석하고자 한다. 개인 간의 소통이 그 어느 때보다 중시되고 있고, 코로나 19로 인해 대면 보다는 비대면을 통한 간접 경험의 기회가 증가하고 있는 지금, SNS는 그 어느 때보다 강한 영향력을 가지고 있다. 즉, 이는 SNS가 그 어느 때보다 사람의 심리와 행동 특성을 잘 반영하는 지표라는 것이다. 어떠한 학자들은 현대 인류의 별칭으로 ‘포노 사피엔스’를 제안하면서 스마트폰, 인터넷, 그리고 SNS의 영향력을 높게 평가한다. 이에, 본 연구는 대상의 실제 드러나는 특징과 온라인상의 특징을 분석하기 위해, 특히, 젊은 세대들이 가장 많이 사용하는 인스타그램과 정보 분류 기능인 해시태그에 집중해, 인스타그램 게시물 및 해당 게시물의 해시태그를 분석하고자 한다.

II. 이론적 배경

1. 인스타그램 특성

인스타그램(Instagram)은 SNS 매체 중 하나로, SNS의 특징을 가장 잘 보여주는 매체 중 하나로 평가되며 다양한 SNS 플랫폼 중 젊은 세대의 사용률 및 만족도가 높은 플랫폼으로도 알려져 있다. 인스타그램은 한 사람 또는 한 집단의 정체성을 가진 계정이 자신의 개성을 보여주고 상호 간의 소통을 하면서 기능한다. 독립된 계정은 스스로 게시물과 ‘스토리(Story)’를 게시해 다수에게 자신의 일상을 공유할 수 있으며, 계정 간의 직접적인 연결은 ‘다이렉트 메시지(Direct Message)’ 기능을 통해 일대일로 대화를 지원하고 있다.

이러한 인스타그램의 기능은 SNS가 가지는 특성을 가장 잘 드러낸다. 사진 또는 짧은 동영상과 글로 구성된 게시물은 삭제가 없다면 영구적으로 게시가 되어 해당 계정이 다른 사람들에게 지속해서 드러내고 싶은 특성을 보여주고, 해당 게시물 아래에서 댓글 기능을 통해 계정 간의 공개적인 소통을 가능하게 한다. 24시간 동안 공개되는 스토리는 해당 계정이 공유하고 싶은 일상을 소소하게 공유하게 해 계정 간의 상호에 대한 이해 및 공감을 더욱 끈끈하게 가능하게 하며 답장을 통해 다이렉트 메시지로 연결되어 지속적인 개인과 개인 사이의 연결을 가능하게 한다. 특히, 스토리 기능에서는 감정을 표현할 수 있는 수단을 다양하게 제공하게 해주어, 사람과 사람 사이를 연결하는데 가장 중요한 감정적인 면을 온라인상에서도 공유할 수 있게 해 진정한 ‘소셜 네트워크 서비스’로서의 역할을 드러낸다. 이러한 기능들은 인스타그램의 계정들이 서로 ‘팔로우(follow)’하므로 가능한 일인데, 계정들이 능동적으로 자신이 원하는 계정을 팔로우함으로써 그 계정에서 공유하는 게시물과 스토리들을 자동으로, 또 지속해서 받아 소통할 수 있다.

인스타그램 게시물의 특성에서 제외할 수 없는 것이 ‘해시태그(Hashtag)’이다. 해시태그는 인스타그램 계정이 주로 게시물을 공유하면서 짧게 뒤에 덧붙이는 글귀로, # 표시 이후에 문자열을 통해 설정된다. 해시태그를 통해 각 계정은 해당 게시물과 연관된 특징을 드러내, 자신의 경험이나 장소, 소속감, 감정 등을 드러낸다. 해시태그를 붙임으로써 계정은 게시물에 특정한 정보를 붙이는 것과 다름이 없다. 이렇게 동일한 해시태그로 묶인 게시물의 집단은 인스타그램 내의 검색 기능을 통해 해시태그 자체를 검색함으로써 찾을 수 있으며, 같은 화면에 동시에 뜨면서 해시태그에 해당하는 정보를 검색자가 간접적으로 경

협하는 데 일조한다.

해시태그의 분류 유형은 연구마다 다르나, 평소 계정들에서 게시물 특징을 드러내기 위해 사용하며 다른 인스타그램 이용자들도 주로 사용하는 유형은 크게 정보 공유형, 콘텐츠 노출형, 감정 공유형으로 나눌 수 있다. 정보 공유형 해시태그는 특정 정보를 공유해 해당 장소, 지역, 경험에 대한 게시물임을 드러낸다. 콘텐츠 노출형 해시태그는 큰 분류 범주의 해시태그에 해당하며, 해시태그를 통해 자신의 게시물과 더 나아가 자신의 계정이 타 계정에 노출되는 것을 유도하기 위함이다. 감정 공유형 해시태그는 감정을 직접 드러내 감정적 공유와 계정 간의 소통을 활성화한다. 이렇게 여러 유형으로 분류할 수 있는 해시태그들은 인스타그램 상에서 정보 및 게시물의 분류를 이용자가 능동적으로 할 수 있게 하여 이용자들이 자신의 개성을 드러내게 해 만족감을 높여줄 뿐 아니라, 정보 분류 작업을 용이하게 해 인스타그램 상에서 해당 정보를 찾는 이용자가 실제 경험이 담긴 생생한 정보를 찾을 수 있게 한다는 점에서 인스타그램의 SNS 기능에 큰 특징이며 다른 플랫폼과 비교해 이점을 제공한다.

이렇게 해시태그는 정보 기능적으로도 큰 특징을 가지고 있으며, 그 특징은 다음과 같이 네 가지로 들 수 있다: 용이성, 유용성, 신뢰성, 그리고 습관성이다. 용이성이란, 적은 노력만으로도 정보를 쉽게 찾을 수 있다는 뜻으로, 앞서 상술했듯이 해시태그는 정보를 찾고자 하는 이용자가 검색만 하면 해당 정보와 연관된 게시물을 찾을 수 있으므로 사용자 입장에서 편리하다. 또, 유용성이란, 데이터의 군집화에 크게 이바지한다는 뜻으로, Rosa(2011)의 연구 등 본 연구 이전의 여러 연구 결과에 따르면 해시태그가 계정 이용자 스스로 자신의 게시물을 분류하기 때문에 능동적이고 실질적인 분류를 하는 데 도움이 되어, 그 유용성이 높다고 할 수 있다. 다음으로, 신뢰성이란 해시태그를 이용한 정보의 분류가 진정성이 있고 믿을 만하다는 뜻으로, 앞서 설명했듯 이용자들이 스스로 분류하기 때문에 진정성 있는 경험이 해시태그 내에 내재되어 있으며, 인스타그램 이용자들도 해시태그의 생성 과정을 이해하고 있기 때문에 이를 객관적으로 판단하고 받아들인다. 그리고 습관성이란, 정보의 이용자들이 지속해서, 또 무의식적으로 기능을 이용해 정보를 군집 및 분류화하며, 또 같은 과정으로 노출이 된다는 뜻으로, SNS의 지속성 및 중독성과 결부해 해시태그는 그 습관성이 강하다고 할 수 있다. 결론적으로 이러한 특성을 가지는 해시태그에서 주목해야 할 것은, 인스타그램의 계정들이 스스로 자신의 게시물을 분류한다는 것이다. 이것이 인스타그램의 직관적인 검색 기능과 합쳐져 용이하며 유용성 있고 신뢰성을 가지는 해시태그의 특징을 만들어내는 것이다.

이러한 특징을 가지는 해시태그는, 온라인 플랫폼의 주 이용층인 20·30세대가 경제활동의 한 축이 된 지금, 방문 장소에 대한 특징적인 정보를 찾는 데에도 큰 역할을 하고 있다. 앞서 설명한 특징들을 바탕으로, 젊은 세대는 특정 장소 및 지역에 대한 정보를 찾기 위해 해시태그를 이용하고 있는데, 블로그와 같은 긴 글을 읽을 필요도 없이 짧은 시간 내에 단편적으로 시각적인 이미지와 더불어 정보들을 제공해주기 때문에 간접 경험으로서 방문 여부를 결정하는 데 역할을 한다. 이용일(2018), 추민수(2021), 임현숙(2019) 등의 연구에 따르면 해시태그의 특징으로 말미암아 해시태그를 통해 SNS 이용자들이 적극적으로 방문 장소에 대한 특징을 조사하고, 이에 대해 신뢰를 한다는 것을 알 수 있으며, 이에 방문 선택 요소로서 작용한다고 할 수 있다.

2. 웹 스크래핑 및 시각화

본 연구는 웹 스크래핑을 통한 빅데이터 분석을 이용해 결과를 도출하는 것을 목적으로 두고 있다. 빅데이터는 연구대상을 특정 집단으로 제한하는 것이 아니기에 사회 전반의 방대한 양의 온라인 데이터를 수집하고 해당 데이터를 분석 및 활용할 수 있는 프로그램을 사용하여 사회 전반의 인식과 시대의 현상을 알아보기 적합한 도구이다.

빅데이터 분석방법에는 텍스트 마이닝, 감성 분석, 평판 분석 등이 있는데, 이때 텍스트 마이닝은 빅데이터 분석방법 중 핵심기술이며, 최근 해당 방법을 이용한 다양한 연구가 진행되고 있다. 이에 본 연구에서는 텍스트마이닝 분석방법에 해당하는 ‘해시태그 분석’을 은평한옥마을을 관련 인스타그램 게시물에 해시태그를 분석하는 연구방법으로 채택했다. 실제 관련 선행연구인 ‘빅데이터 분석방법을 적용한 인스타그램 뷰티관련 해시태그 게시물 분석에 관한 융합 연구’에서는 인스타그램에서 뷰티, 메이크업 관련 게시물에 해시태그를 추출하고 이 해시태그들의 키워드를 분석하여 메이크업과 화장품에 대한 소비자들의 인식 분석 연구를 진행하였다. 그뿐만 아니라, 지자체의 관광명소 관련 SNS의 해시태그를 추출하여 관광산업의 발전 방안을 제시하였으며, 트윗을 분석하여 영화관에서 상영되는 영화에 대한 관객들의 선호도를 분석한 사례도 존재한다. 이처럼 트렌드 및 소비자의 인식을 파악하기 위해 웹 스크래핑을 통한 SNS 텍스트마이닝 분석방법이 많은 분야에서 이용되고 있다.

또한, SNS에서 데이터를 추출하여 가공 및 분석을 하고 데이터를 시각화하여 결과를 시각화하는 방법도 중요시되고 있다. 이에 본 연구에서는 데이터 시각화 과정으로 ‘워드클라우드’를 사용했다. 워드클라우드는 웹 스크래핑을 통해 추출한 해시태그를 분석하여 해시태그별 노출 빈도수를 고려하여 시각적으로 나열한 것으로서 기본적인 시각화 방법에 해당한다. 데이터의 빈도수나 중요도에 따라 글자의 굵기와 색상에 변화를 주는 방법이다. 본 연구에서는 해시태그를 추출하고 이를 가공하여 중요 키워드 선정 및 키워드별 중요도를 수치화하였으며 이를 워드클라우드로 데이터를 시각화하였다.

Ⅲ. 연구 방법

1. 연구내용 및 방법

인스타그램이 사람들의 인식을 반영한다는 것은 자명하게 알려져 있다. 이에 기반하여 본 연구에서는 인스타그램을 활용하여 해시태그를 분석하고 이를 통하여 은평한옥마을에 존재하는 카페, 주위 음식점, 그리고 방문요인을 중심으로 은평한옥마을의 특징을 분석하고자 한다.

관련 선행연구 조사 결과, 이와 유사한 연구에서 주로 사용되는 Python 모듈로는 Requests, Selenium, BeautifulSoup 등이 있다. 본 연구에서는 SNS 중 인스타그램 스크래핑 목적에 맞게, 그리고 연구의 효율성을 높이고자 Instaloader를 사용하여 Python 웹 스크래핑을 진행했다. 따라서 인스타그램에서 “은평한옥마을”과 “은평한옥마을카페” 해시태그가 포함된 인스타그램 게시물을 Instaloader를 이용해 웹 스크래핑하여 그 게시물에 포함된 다른 해시태그, 게시 시점, 좋아요 수, 조회수, 댓글 수 등등의 데이터를 추

동일한 방식으로 분석을 진행하였다. ‘은평한옥마을카페’ 해시태그 검색결과 인기가 있다고 판단할 수 있는 은평한옥마을 주위 카페들로는 1인1잔(12위, 1230개), 북한산제빵소(56위, 298개), 로맨티카(70위, 207개), 롱브레드(76위, 187개), 스페이스헬레나(95위, 149개), 카페 불가심(136위, 94개) 등이 있었다. 이 중에서 1인1잔, 스페이스헬레나, 카페 불가심은 한옥 카페인 점이 특징적이다. ‘은평한옥마을’ 해시태그와 ‘은평한옥마을카페’ 해시태그에서 카페들의 순위를 보아 카페 로맨티카의 순위에 변동이 있지만 대부분 비슷한 경향성을 보였으며, 1인1잔이 각각 6500개, 1230개의 해시태그로 압도적인 비율을 차지하고 있다.

카페 (해시태그 수)	
1인 1잔	6500
북한산 제빵소	2498
롱브레드	766
스페이스헬레나	434
카페 불가심	399
로맨티카	333
투썸플레이스	177
카페베네	94

<표 1> ‘은평한옥마을’으로 추출한 카페와 관련된 해시태그

카페 (해시태그 수)	
1인 1잔	1230
북한산 제빵소	298
로맨티카	207
롱브레드	187
스페이스헬레나	149
카페 불가심	94
투썸플레이스	22

<표 2> ‘은평한옥마을카페’로 추출한 카페와 관련된 해시태그

2) 음식점 및 베이커리

‘은평한옥마을’ 해시태그를 검색한 경우 은평구맛집(20위, 1321개), 멕스타그램(27위, 1084개) 등으로 카페뿐만 아니라 음식점을 방문하기 위해 은평한옥마을을 찾은 사람들이 있을 것으로 생각하였고 관련 해시태그들을 찾아본 결과 아래 표와 같이 1인1상(14위, 1688개), 오스테리아피우(50위, 796개), 롱브레드(53위, 766개), 다운브레드(54위, 749개) 등을 확인할 수 있었다.

음식점 및 베이커리 (해시태그 수)	
1인 1상	1688
오스테리아피우	796
롱브레드	766
다운브레드	749

<표 3> ‘은평한옥마을’로 추출한 음식점 및 베이커리와 관련된 해시태그

3) 풍경

‘은평한옥마을’ 해시태그를 검색한 경우 북한산(4위, 5047개), 진관사(10위, 2600개), 북한산 둘레길(23위, 1304개) 등과 같이 근처 풍경과 관련된 해시태그들을 순위권에서 쉽게 확인할 수 있었다.

풍경 관련 (해시태그 수)	
북한산	5047
진관사	2498
북한산 둘레길	1304

<표 4> ‘은평한옥마을’로 추출한 풍경과 관련된 해시태그

4) 방문요인

‘은평한옥마을’을 검색했을 때 함께 나타나는 해시태그 중 은평한옥마을 방문요인으로 사료되는 해시태그들을 분류하면 일상(3위, 5862개), 한옥(7위, 3189개), 카페(8위, 2962개), 사진(18위, 1354개), 주말(21위, 1315개), 먹스타그램(27위, 1084개), 데이트(31위, 998개), 여행(32위, 989개), 뷰맛집(38위, 897개) 등이 있었다.

주목해볼만한 해시태그들 (해시태그 수)	
일상	5862
한옥	3189
카페	2962
사진	1354
주말	1315
먹스타그램	1084
데이트	998
여행	989
selfie	911
뷰맛집	897
아람채	299

<표 5> ‘은평한옥마을’로 추출한 방문 동기와 관련된 해시태그

5) 은평한옥마을에서 벗어난 지역: 경기도 고양시(일산)

‘은평한옥마을’를 검색했을 때의 해시태그 키워드 중 17위(1382개, 일산카페), 24위(1271개, 일산호수공원), 25위(1231개, 벨라시타) 등의 순위를 차지했던 타 지역 해시태그는 ‘은평한옥마을카페’를 검색했을 때 더욱 뚜렷한 양상을 보였다. 경기도 고양시 관련 해시태그는 2위와 6~9위, 13~17위, 20위 등을 차지했으며, 이 밖의 지역으로는 김포카페, 파주카페가 공동 21위(1029개)를 차지했다.

고양시 관련 해시태그들 (해시태그 수)	
일산카페	2167
일산가로수길카페	1329
일산카페추천	1312
일산호수공원	1311
일산디저트	1299
벨라시타	1203
일산맛집	1149
서오릉카페	1136
행주산성카페	1126
일산카페투어	1100
애니골카페	1065

<표 6> ‘은평한옥마을’로 추출한 고양시와 관련된 해시태그

2. 결과 해석

본 연구에서는 전체 게시물을 바탕으로 은평한옥마을의 활성화 정도, 콘텐츠의 종류와 해시태그 개수를 바탕으로 은평한옥마을의 특징을 분석하고자 한다. 그리고 은평한옥마을의 특징을 부각하기 위하여 북촌한옥마을을 비교 대상을 선정하였다.

(1) 전체 게시물

전체 게시물의 관점에서 ‘은평한옥마을’ 태그를 한 게시물은 47227개, ‘은평한옥마을카페’를 태그한 게시물은 2915개가 존재한다. 하지만 북촌 한옥마을의 경우 ‘북촌한옥마을’ 태그를 한 게시물은 315583개로 은평한옥마을의 약 6배에 이르렀다. 이처럼 은평한옥마을이 북촌한옥마을과 비교하면 활성화의 정도가 비교적 작다는 것을 알 수 있다.

(2) 콘텐츠의 종류

은평한옥마을의 콘텐츠는 4.1에서 확인할 수 있듯이 크게 카페, 음식점, 풍경, 방문 동기, 그리고 은평한옥마을과 관련이 없는 태그 즉, 5가지로 나눌 수 있다.

첫 번째, 카페에 대해서는 한 카페에 의존도가 높은 것을 알 수 있다. 1인 1잔이 해시태그가 6500개로 압도적인 1위이며, 2위인 북한산 제빵소의 경우 2498개, 3위인 롱브레드가 766개로 순위가 낮아질수록 관련 해시태그의 개수가 현저하게 줄어든다. 따라서 이를 미루어 보아 은평한옥마을은 ‘1인1잔’이라는 카페에 의존도가 매우 높고, 다른 카페들은 인스타그램에서 홍보나 언급이 1인1잔에 비해 매우 적다는 것을 알 수 있다.

두 번째, 음식점 및 베이커리와 관련해서는 타 지역에 비해 업소 수가 절대적으로 적음을 확인할 수 있었다. 1인1상, 오스테리아피우, 다운브레드, 롱브레드로 총 4개의 주요 해시태그밖에 없으며, 이와 관련한 게시물 수도 매우 적기 때문이다. 이를 바탕으로 은평한옥마을에서는 음식점 및 베이커리의 선택지가 매우 좁고, 사람들이 즐겨 찾는 맛집이 없음을 알 수 있다.

세 번째, 풍경에 대해서는 북한산, 진관사가 상위권을 차지했으며 사람들이 풍경을 보기 위해 은평 한옥마을을 찾음을 알 수 있었다. 해시태그의 ‘풍경’ 부문에 있는 ‘북한산’ 해시태그는 5047번 언급되어 전체 해시태그 순위 중 4위를 차지하였다. 또한, 해시태그의 ‘방문요인’ 부문에서는 ‘사진’ 해시태그가 1354번, ‘뷰맛집’ 해시태그가 897번 언급된 것으로 보아 은평 한옥마을 방문자들이 북한산 풍경을 감상하는 것과 북한산을 배경으로 사진을 즐겨 찍는다는 것을 알 수 있었다. 종로의 도시풍경을 배경으로 하는 북촌한옥마을과 달리 도심 속에서 흔히 볼 수 없는 북한산이라는 자연적인 풍경을 배경으로 하는 은평한옥마을은 한옥과 조화로움을 형성한다. 이러한 부분에 있어 북한산은 은평한옥마을만의 장점이라고 할 수 있다.

네 번째, 은평한옥마을을 방문하는 이유(방문요인)로는 매우 다양하게 존재함을 알 수 있다. 그리고 이때 일상이 1위이며 5862개의 게시물을 포함하고 있는 것으로 보아 주로 인스타 인플루언서나 무직인 사람들이 방문함을 알 수 있다. 주로 이러한 사람들이 카페(3위, 2962개)에 가거나 먹스타그램(6위, 1084개)을 보여주기 위해 은평한옥마을에 방문하였음을 알 수 있다. 그 외에도 주말(5위, 1315개)에 시간이 나서 방문하거나 사진(4위, 1354개)을 찍기 위해 혹은 데이트(7위, 998개)를 하기 위해 방문하였음을 알 수 있다.

마지막으로, 은평한옥마을에서 벗어난 지역에 대한 해시태그가 다량으로 있는 것에는 여러 가지 가능성이 존재한다. 하지만 은평 한옥마을이 경기도 고양시와 지리적으로 근접한 위치에 있다는 점이 은평한옥마을을 해시태그로 하여 검색했을 때 고양시와 관련된 많은 태그가 등장하는 이유로 사료된다.

V. 결론 및 제언

본 연구는 은평한옥마을에 관련된 인스타그램 해시태그 분석을 통해 은평한옥마을 방문자들의 방문요인과 은평한옥마을의 콘텐츠의 종류와 수를 조사하였다.

본 연구에서 조사한 인스타그램 해시태그의 경향으로 보아 이곳을 방문하는 사람들은 카페에서 풍경을 바라보며 여유를 즐기기 위해 은평 한옥마을을 찾는 경우가 많음을 알 수 있으며 카페의 경우 ‘1인 1잔’의 의존도가 높음을 알 수 있었다. 이러한 이유는 1인 1잔의 많은 홍보성 게시글에 영향을 받았음을 알 수 있다. 반대로, 이는 전반적으로 은평한옥마을의 콘텐츠 자체가 다양하지 않음을 의미하며 이는 음식점에서 특히 두드러진다. 이러한 은평한옥마을의 특성은 은평한옥마을이 관광지가 아닌 주거지를 목적으로 개발된 곳이라는 점에서 기인한다고 볼 수 있다. 마찬가지로 이는 은평 한옥마을이 북촌 한옥마을보다 주변 인구 대비 활성화 정도가 낮은 이유라고 할 수 있다.

VI. 참고 문헌

- 1) 김예지, 김학선.(2019).빅데이터 분석을 통한 디저트 카페 개념에 대한 탐색적 연구.한국조리학회지,25(5),125-135.
- 2) 남미우.(2019).빅데이터 분석 방법을 적용한 인스타그램 뷰티관련 해시태그 게시물 분석에 관한 융합 연구.한국과학예술융합학회,37(4),93-106.
- 3) 이애리, 이주원.(2018).소셜 빅 데이터를 이용한 상권 확장 트렌드 및 소비 트렌드 분석.e-비즈니스연구,19(6),401-413.

VII. 부록

1. Instaloader를 활용하여 웹 스크래핑 (제작 협조: 기술팀 박정연 A)

① Python 프로그램

```
import instaloader
from datetime import datetime
from itertools import dropwhile, takewhile

bot = instaloader.Instaloader(
    download_pictures=False,
    download_videos=False,
    download_comments=False,
    save_metadata=True,
    request_timeout=100,
    compress_json=False,
)

bot.login("(아이디)", "(비밀번호)")
bot.download_hashtag('은평한옥마을', max_count=80000, fast_update=False,
profile_pic=False)
```

② Windows PowerShell에서 실행

```
PS C:\Users\seoki\Desktop\지역연계\지역연계 webscraping> instaloader --login (아이디) -p (비번) "#은평한옥마을" --no-pictures --no-videos --no-video-thumbnails --no-captions --no-compress-json
```

2. .json 파일에 대한 파싱을 위한 Python 파일 (제작 협조: 기술팀 이현서 A)

```
import json
import datetime
from datetime import date
from datetime import datetime
import re
import os
import csv

dir = "C:\\Users\\seoki\\Desktop\\#은평한옥마을카페"
files = os.listdir(dir) #파일 (리스트)

jsons = []
txts = []
for i in files:
    if(i[-1]=="n"):
        jsons.append(i)
    else:
        txts.append(i)
jsons.pop(0)
"""
print(jsons) #json 파일들
print(txts) #txt 파일들
"""
print(len(jsons))
def parsejson(file): #file 매개변수에 넣은 이름을 가진 json파일에서 좋아요수, 날짜, 댓
글수 반환
    with open(dir+"\\"+file,"r") as f: #2014-11-05_08-43-08.UTC.json
        data= json.load(f)
        likes = data["node"]["edge_liked_by"]["count"] #좋아요 수
        #print(likes)
        timestamp = data["node"]["taken_at_timestamp"]
        date_obj = date.fromtimestamp(timestamp) #업로드 날짜
        #print(date_obj)

        comments = data["node"]["edge_media_to_comment"]["count"] #댓글수 (?)
        #print(comments)
        return likes, date_obj, comments
```

```

def parsehashtag(file): #file 매개변수에 넣은 이름을 가진 txt 파일에서 해시태그 반환
    with open(dir+"\\"+file,"r",encoding="utf-8") as f:
#2014-11-05_08-43-08.UTC.txt
        data = f.read()
        startpoints=[]
        endpoints=[]
        start=None
        end=None
        for i in range(0,len(data)):
            if(data[i]=="#"):
                if(start==None):
                    start =i
                    startpoints.append(i)
                else:
                    end =i
                    endpoints.append(end)
                    start=i
                    startpoints.append(i)

            elif(data[i]==" " or data[i]=="\n"):
                if(start==None):
                    continue
                else:
                    end=i
                    endpoints.append(end)
                    start =None

        print(startpoints,"\n",endpoints)
        list=[]
        for i in range(0,len(startpoints)):
            list.append(data[startpoints[i]:endpoints[i]])
        for i in range(len(list)):
            list[i]=list[i].replace("#","")
            list[i]=list[i].replace("\n","")

        return list

```



```

with open(dir+"\\"+"insta.csv","w") as f: #dir+"\\"+
    writer = csv.writer(f)
    writer.writerow(["좋아요 수","업로드 날짜","댓글 수","해시태그"])
    hashdict = {}
    for i in jsons: #모든 json 파일 순회
        try:
            likes , upload_date, comments = parsejson(i)
            txt= i[:len(i)-4]+"txt"
            hashtags = parsehashtag(txt)
            hashtags_str = ",".join(hashtags)
            for i in hashtags:
                if i in hashdict.keys():
                    hashdict[i]+=1
                else:
                    hashdict[i]=1
            print(likes,upload_date,comments,hashtags)
            writer.writerow([likes,str(upload_date),comments,hashtags_str])
        except:
            continue
    print(hashdict)

with open(dir+"\\"+"hashtags.csv","w") as f: #csv 파일 생성
    writer = csv.writer(f)
    writer.writerow(["해시태그","수"])
    for key in hashdict.keys():
        try:
            writer.writerow([key,int(hashdict[key])])
        except:
            continue

```