

과제03: 데이터와 텍스트

언어 데이터 처리 2022-11-27 장태준 [컴퓨터 공학부 2017-17018]

<https://github.com/jtjun/NLP/blob/main/hw03/hw03.md>

1. 개방형 표준

A. 개방형 표준(open standard)

개방형 표준(open standard), 공개 표준은 기술 표준이 문서가 공개되어 있으며 사용이 자유로운 경우를 가리키는 용어이다.

개방형 표준은 상황에 따라 다른 정의가 존재하며, 가령 여러 표준 단체나 국가마다 서로 다른 정의를 가진다. 일반적으로 포함되는 개념은 표준을 사용하는 데에 있어서 로열티를 지불할 필요가 없다는 점이다. 이때 일부 경우는 기술 표준을 사용하는 데에 특히 라이선스가 필요하며, 합리적이고 비차별적인(RAND) 라이선스를 허용하는 하에 해당 표준을 개방형 표준으로 인정하기도 하다.

따라서 개방형 표준은 통합된 환경을 제공할 수 있어, 소프트웨어 - 기업 - 나라 간의 상호작용에서 발생할 수 있는 비용을 감소시키며, 종속성 문제를 해결한다. 실시간으로 국가간 데이터를 교환하는 글로벌 시대에 이러한 개방형 표준은 핵심적인 역할을 하고 있다.

B. 개방형 표준 중 파일 형식

- ODF (OpenDocument Format) : ISO 문서 표준 형식
- OOXML (Office Open XML) : Microsoft에서 공개한 ISO 표준 문서 형식
- PDF (Portable Document Format) : Adobe에서 공개한 서식 있는 문서를 위한 ISO 표준 형식
- PNG (Portable Network Graphics) : 비트맵 이미지 ISO 표준 형식
- HTML (Hypertext Markup Language) : (W3C)
- SVG (Scalable Vector Graphics) : 벡터 그래픽을 위한 오픈 포맷 (W3C)
- Ogg : 멀티미디어를 위한 오픈 포맷

C. ODF 와 OOXML 그리고 OWPML

ODF는 OASIS 컨소시엄 산하 기술위원회(Technical Committee; TC)가 제정하였다. ODF-TC에는 많은 업체들과 개인이 동등한 자격으로 참가하고 있으며, 많은 오피스 프로그램과 관련 문서 시스템의 개발자들이 표준화 작업을 진행하고 있다.

이에 반해 OOXML은 마이크로소프트가 추진하였으며 2006년과 2008년 두 차례에 걸쳐 Ecma 인터내셔널에 유럽 표준(ECMA-376)으로, 2008년에 ISO/IEC에 국제 표준(ISO/IEC 29500)으로 승인되었다.

한 편, 두 표준 모두 머신리더블한 포맷이기에, 별도의 데이터 처리 과정을 거치거나 특수한 프로그램을 사용하지 않고 문서에 담긴 데이터를 다룰 수 있다. 그러나 우리 나라에서 널리 쓰이는 아래아한글의 경우, 머신리더블하지 않은 hwp 형식을 사용하고 있었다.

디지털전환이 가속화됨에 따라 전자문서 데이터의 중요도가 더욱 높아지고 있는 만큼, 아래아한글을 머신리더블한 HWPE로 본격 변환함으로써 전자문서 국제 표준화를 선도하고 활용도를 더욱 넓혀 나가기 위해 OWPML을 개발하여 국가표준으로 등록하였다.

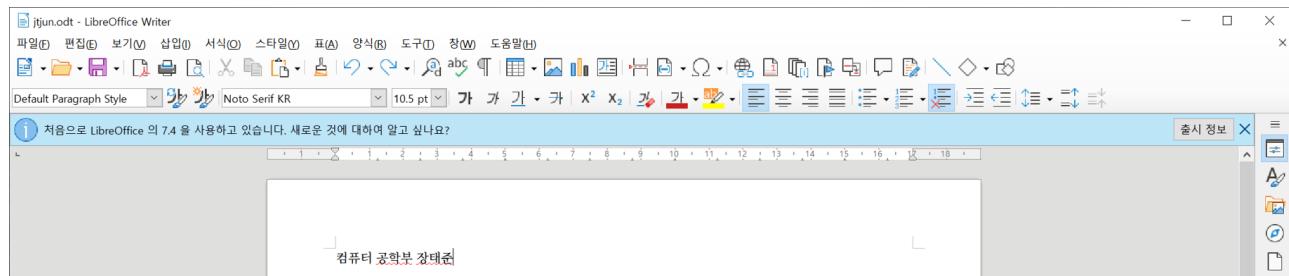
2. 형식별 실습

각 형식이 어떻게 구성되어 있는지 살펴보고 텍스트로 된 데이터로서 사용하는 실습

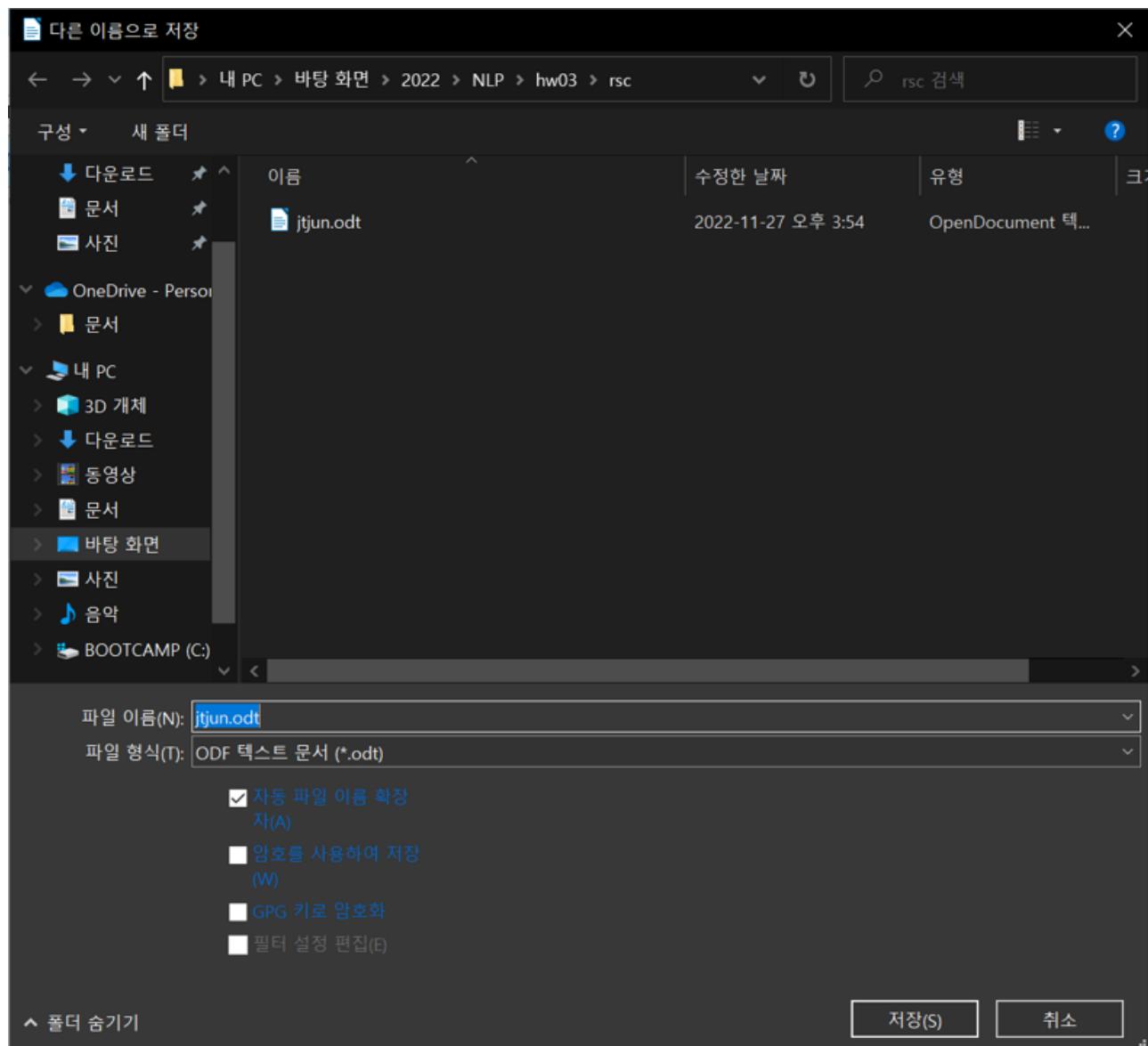
수업 실습 영상강의를 참조, 실습 과정을 스크린샷을 이용하여 보여주고 텍스트로 된 데이터의 구조를 코드블럭으로 예시하고 설명과 의견을 포함하여 보고서를 작성

A. ODF 중 ODT 형식 (워드프로세서)

1. LibreOffice로 ODT 파일 만들기



'컴퓨터 공학부 장태준'이라는 내용으로 새로운 파일을 만든다.



'.odt'형식으로 파일을 저장한다.

2. VS Code에서 확인하기

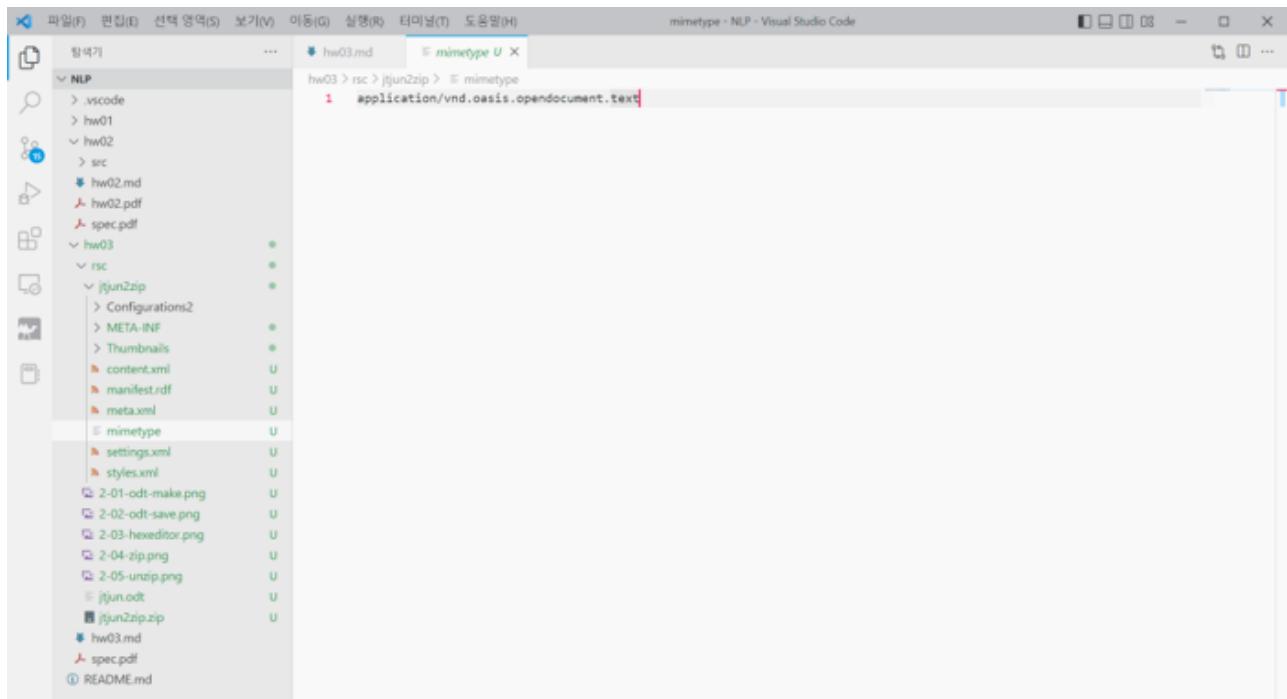
The screenshot shows the Visual Studio Code interface with the file structure of `hw03.rsc`. The `jtjun.odt` file is open in the editor. The `Decoded Text` tab displays the raw hex data as interpreted text, showing various file headers and metadata. The `Data Inspector` tab on the right provides detailed information about the data types and their memory addresses.

먼저, hex-editor로 파일을 확인해본다.

3. 압축해제

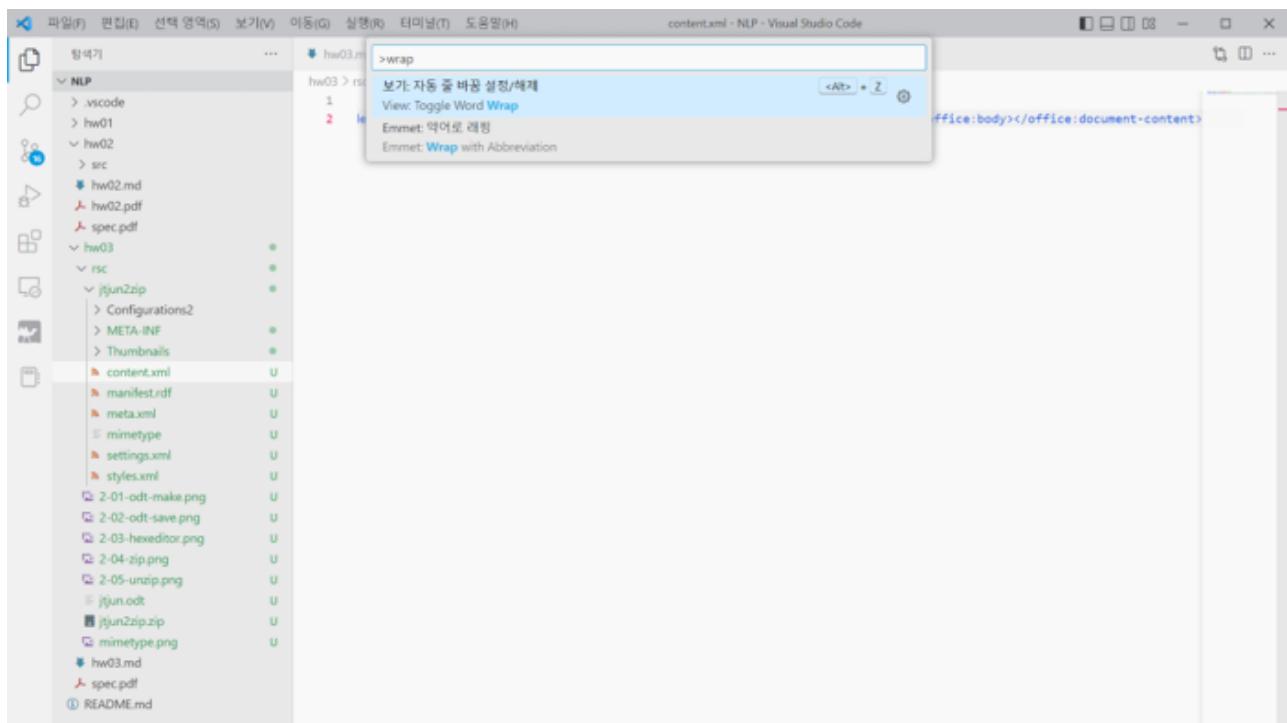
The screenshot shows the WinRAR interface with the file `jtjun2zip.zip` open. The contents pane displays a list of files extracted from the archive, including `Configurations2`, `META-INF`, `Thumbnails`, `content.xml`, `manifest.rdf`, `meta.xml`, `mimetype`, `settings.xml`, and `styles.xml`.

확장자를 zip으로 변경한 뒤 압축해제 한다.

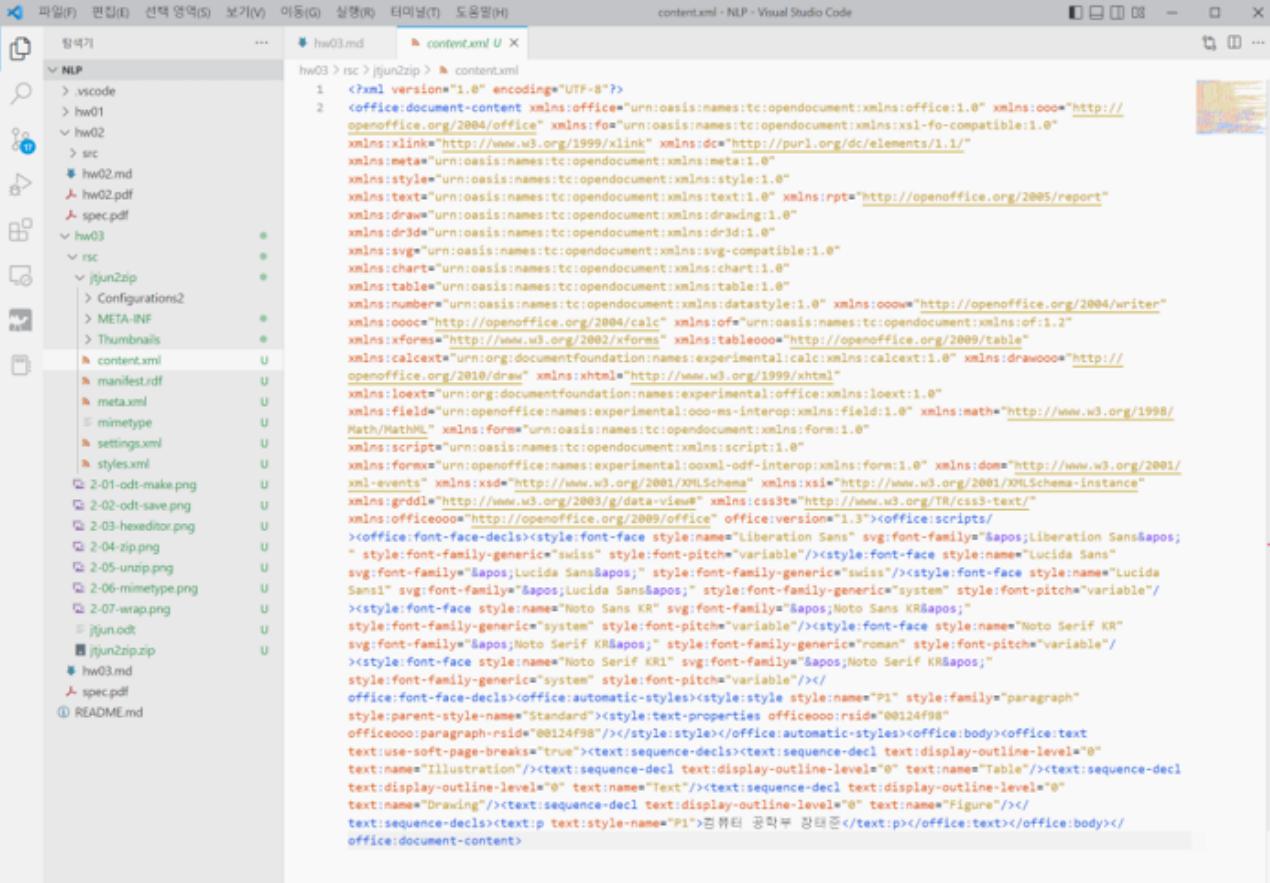


압축해제된 폴더를 VS Code로 확인해본다. 위 캡쳐는 mimetype 파일이다.

4. content.xml 확인



확장 기능을 통해 wrap 형식으로 content.xml 파일을 확인해본다.

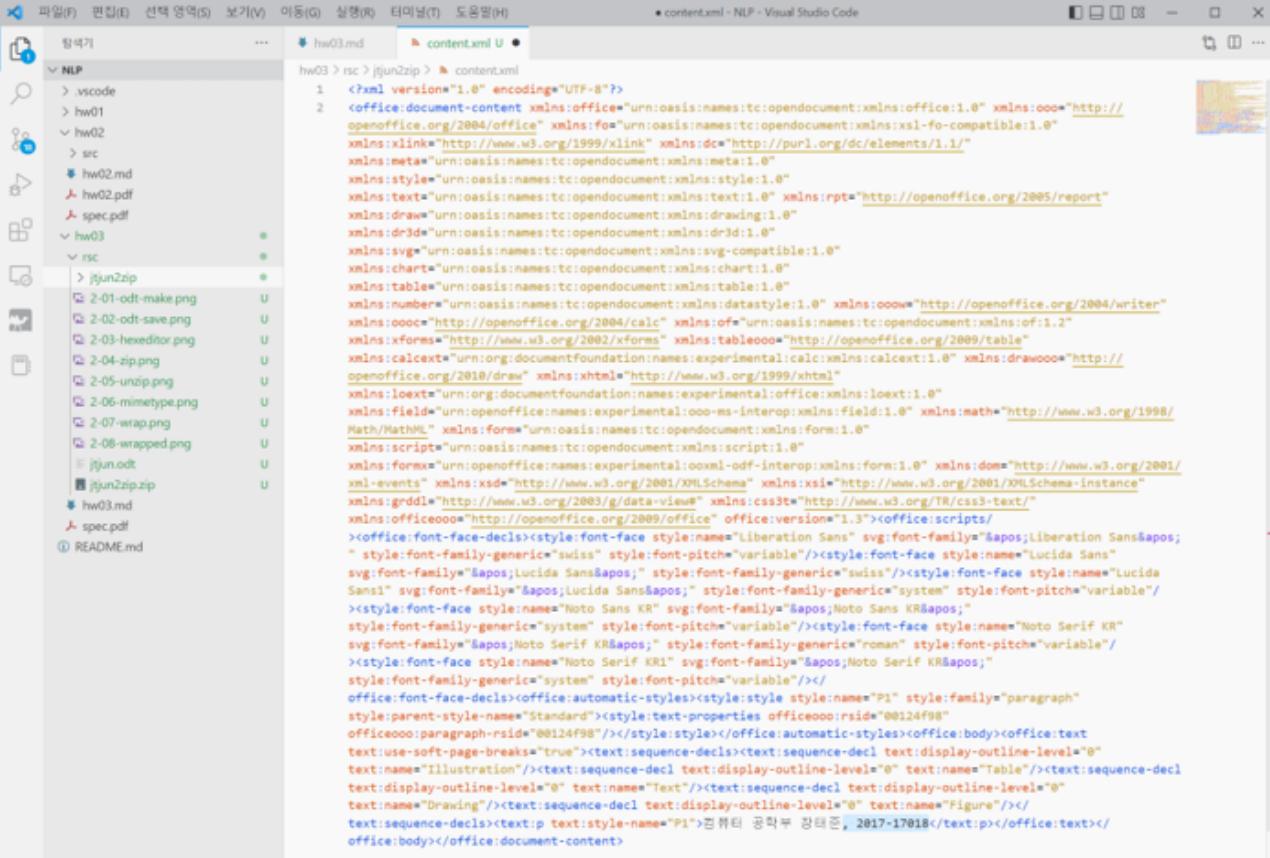


```

<?xml version="1.0" encoding="UTF-8"?>
<office:document-content xmlns:office="urn: oasis:names:tc:opendocument:xmlns:office:1.0" xmlns:ooo="http://openoffice.org/2004/office" xmlns:fo="urn: oasis:names:tc:opendocument:xmlns:xsl-fo-compatible:1.0" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:meta="urn: oasis:names:tc:opendocument:xmlns:meta:1.0" xmlns:style="urn: oasis:names:tc:opendocument:xmlns:style:1.0" xmlns:text="urn: oasis:names:tc:opendocument:xmlns:text:1.0" xmlns:rpt="http://openoffice.org/2005/report" xmlns:draw="urn: oasis:names:tc:opendocument:xmlns:drawing:1.0" xmlns:dr3d="urn: oasis:names:tc:opendocument:xmlns:dr3d:1.0" xmlns:svg="urn: oasis:names:tc:opendocument:xmlns:svg-compatible:1.0" xmlns:chart="urn: oasis:names:tc:opendocument:xmlns:chart:1.0" xmlns:table="urn: oasis:names:tc:opendocument:xmlns:table:1.0" xmlns:number="urn: oasis:names:tc:opendocument:xmlns:datatype:1.0" xmlns:ooow="http://openoffice.org/2004/writer" xmlns:oooc="http://openoffice.org/2004/calc" xmlns:of="urn: oasis:names:tc:opendocument:xmlns:of:1.2" xmlns:xforms="http://www.w3.org/2002/xforms" xmlns:tableooo="http://openoffice.org/2009/table" xmlns:calcext="urn: org: documentfoundation: names: experimental: calc: xmlns: calcext:1.0" xmlns:drawooo="http://openoffice.org/2010/draw" xmlns:xhtml="http://www.w3.org/1999/xhtml" xmlns:loext="urn: org: documentfoundation: names: experimental: office: xmlns: loext:1.0" xmlns:field="urn: openoffice: names: experimental: ooo: ms: interop: xmlns: field:1.0" xmlns:math="http://www.w3.org/1998/Math/MathML" xmlns:form="urn: oasis:names:tc:opendocument:xmlns:form:1.0" xmlns:script="urn: oasis:names:tc:opendocument:xmlns:script:1.0" xmlns:formx="urn: openoffice: names: experimental: ooxml-odf-interop: xmlns: form:1.0" xmlns:dom="http://www.w3.org/2001/xml-events" xmlns:xdm="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:grd1="http://www.w3.org/2003/g/data-view#" xmlns:css3t="http://www.w3.org/TR/css3-text/" xmlns:officeooo="http://openoffice.org/2009/office" office:version="1.3"><office:scripts/><office:font-face-decls><style:font-face style:name="Liberation Sans" svg:font-family="Sapops;Lucida Sans" style:font-family-generic="swiss"/><style:font-face style:name="Lucida Sans" style:font-family="Sapops;Lucida Sans&Sapops;" style:font-family-generic="swiss" style:font-pitch="variable"/><style:font-face style:name="Noto Sans KR" svg:font-family="Kapos;Noto Sans KR&Sapops;" style:font-family-generic="system" style:font-pitch="variable"/><style:font-face style:name="Noto Serif KR" svg:font-family="Sapops;Noto Serif KR&Sapops;" style:font-family-generic="roman" style:font-pitch="variable"/><style:font-face style:name="Noto Serif KR1" svg:font-family="Kapos;Noto Serif KR&Sapops;" style:font-family-generic="system" style:font-pitch="variable"/></style:font-face-decls><office:automatic-styles><style:style style:name="P1" style:family="paragraph" style:parent-style-name="Standard"><style:text-properties officeooo:rsid="#00124f98" officeooo:paragraph-rsid="#00124f98"/></style:style></office:automatic-styles><office:body><office:text text:use-soft-page-breaks="true"><text:sequence-decls><text:sequence-decl text:display-outline-level="0" text:name="Illustration"/><text:sequence-decl text:display-outline-level="0" text:name="Table"/><text:sequence-decl text:display-outline-level="0" text:name="Drawing"/><text:sequence-decl text:display-outline-level="0" text:name="Figure"/><text:sequence-decls><text:p text:style-name="P1">컴퓨터 공학부 장태준</text:p></office:text></office:body></office:document-content>
```

맨 아래 줄에 입력한 내용, '컴퓨터 공학부 장태준'이 들어가있는 것을 확인할 수 있다.

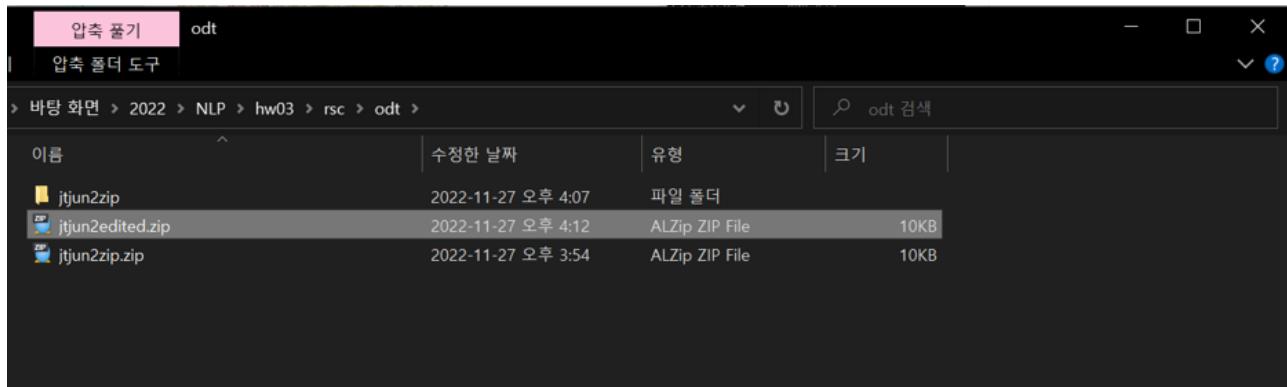
5. xml에서 변경 후 확인하기



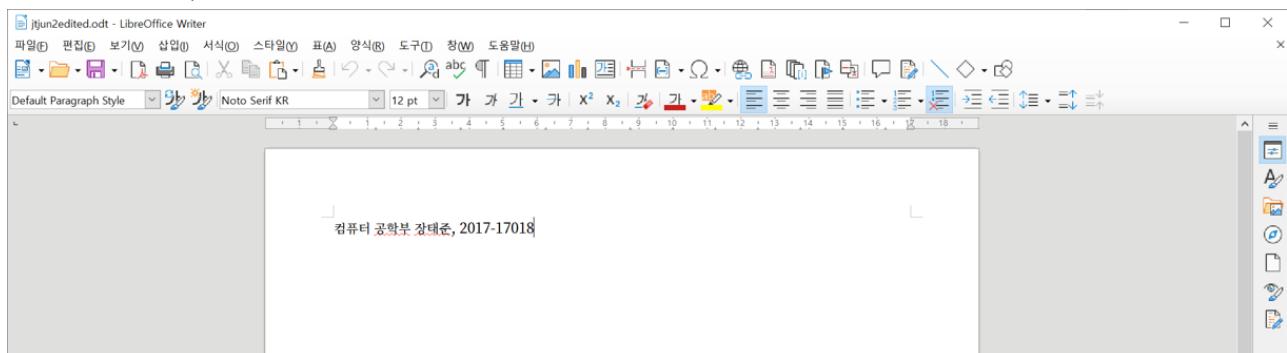
```

<?xml version="1.0" encoding="UTF-8"?>
<office:document-content xmlns:office="urn: oasis:names:tc:opendocument:xmlns:office:1.0" xmlns:ooo="http://openoffice.org/2004/office" xmlns:fo="urn: oasis:names:tc:opendocument:xmlns:xsl-fo-compatible:1.0" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:meta="urn: oasis:names:tc:opendocument:xmlns:meta:1.0" xmlns:style="urn: oasis:names:tc:opendocument:xmlns:style:1.0" xmlns:text="urn: oasis:names:tc:opendocument:xmlns:text:1.0" xmlns:rpt="http://openoffice.org/2005/report" xmlns:draw="urn: oasis:names:tc:opendocument:xmlns:drawing:1.0" xmlns:dr3d="urn: oasis:names:tc:opendocument:xmlns:dr3d:1.0" xmlns:svg="urn: oasis:names:tc:opendocument:xmlns:svg-compatible:1.0" xmlns:chart="urn: oasis:names:tc:opendocument:xmlns:chart:1.0" xmlns:table="urn: oasis:names:tc:opendocument:xmlns:table:1.0" xmlns:number="urn: oasis:names:tc:opendocument:xmlns:datatype:1.0" xmlns:ooow="http://openoffice.org/2004/writer" xmlns:oooc="http://openoffice.org/2004/calc" xmlns:of="urn: oasis:names:tc:opendocument:xmlns:of:1.2" xmlns:xforms="http://www.w3.org/2002/xforms" xmlns:tableooo="http://openoffice.org/2009/table" xmlns:calcext="urn: org: documentfoundation: names: experimental: calc: xmlns: calcext:1.0" xmlns:drawooo="http://openoffice.org/2010/draw" xmlns:xhtml="http://www.w3.org/1999/xhtml" xmlns:loext="urn: org: documentfoundation: names: experimental: office: xmlns: loext:1.0" xmlns:field="urn: openoffice: names: experimental: ooo: ms: interop: xmlns: field:1.0" xmlns:math="http://www.w3.org/1998/Math/MathML" xmlns:form="urn: oasis:names:tc:opendocument:xmlns:form:1.0" xmlns:script="urn: oasis:names:tc:opendocument:xmlns:script:1.0" xmlns:formx="urn: openoffice: names: experimental: ooxml-odf-interop: xmlns: form:1.0" xmlns:dom="http://www.w3.org/2001/xml-events" xmlns:xdm="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:grd1="http://www.w3.org/2003/g/data-view#" xmlns:css3t="http://www.w3.org/TR/css3-text/" xmlns:officeooo="http://openoffice.org/2009/office" office:version="1.3"><office:scripts/><office:font-face-decls><style:font-face style:name="Liberation Sans" svg:font-family="Sapops;Lucida Sans" style:font-family-generic="swiss"/><style:font-face style:name="Lucida Sans" style:font-family="Sapops;Lucida Sans&Sapops;" style:font-family-generic="swiss" style:font-pitch="variable"/><style:font-face style:name="Noto Sans KR" svg:font-family="Kapos;Noto Sans KR&Sapops;" style:font-family-generic="system" style:font-pitch="variable"/><style:font-face style:name="Noto Serif KR" svg:font-family="Sapops;Noto Serif KR&Sapops;" style:font-family-generic="roman" style:font-pitch="variable"/><style:font-face style:name="Noto Serif KR1" svg:font-family="Kapos;Noto Serif KR&Sapops;" style:font-family-generic="system" style:font-pitch="variable"/></style:font-face-decls><office:automatic-styles><style:style style:name="P1" style:family="paragraph" style:parent-style-name="Standard"><style:text-properties officeooo:rsid="#00124f98" officeooo:paragraph-rsid="#00124f98"/></style:style></office:automatic-styles><office:body><office:text text:use-soft-page-breaks="true"><text:sequence-decls><text:sequence-decl text:display-outline-level="0" text:name="Illustration"/><text:sequence-decl text:display-outline-level="0" text:name="Table"/><text:sequence-decl text:display-outline-level="0" text:name="Drawing"/><text:sequence-decl text:display-outline-level="0" text:name="Figure"/><text:sequence-decls><text:p text:style-name="P1">컴퓨터 공학부 장태준, 2017-17018</text:p></office:text></office:body></office:document-content>
```

맨 아래 줄에 내용을 추가하고,

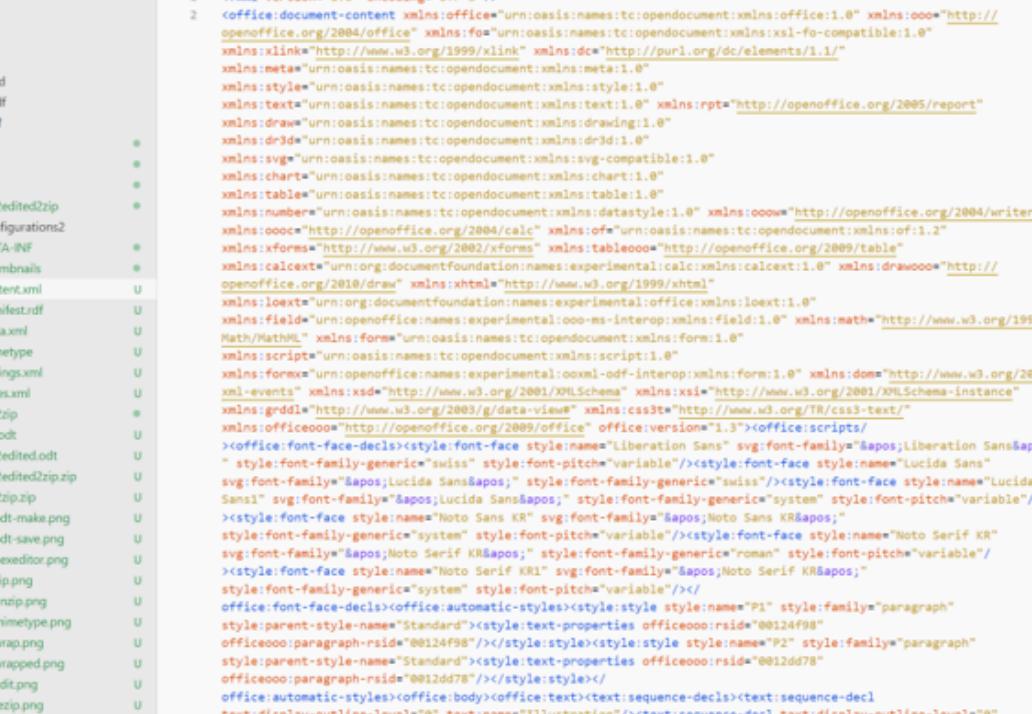


다시 압축한 뒤, 확장자를 .odt로 변경하여 내용을 확인한다.



xml에서 변경한 내용이 제대로 반영되어 있는 것을 확인할 수 있다.

6. odt에서 변경 후 확인하기



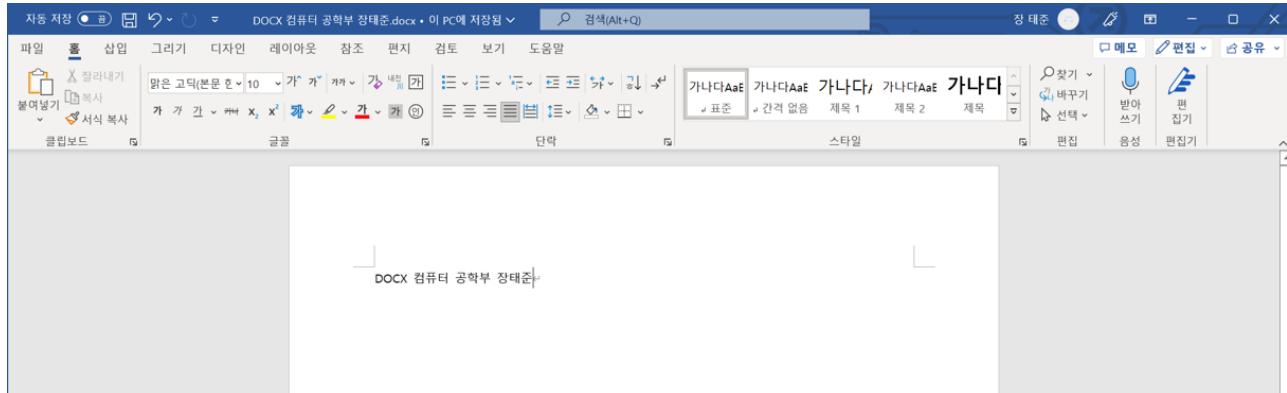
The screenshot shows the Visual Studio Code interface with the following details:

- File Explorer:** On the left, it displays a tree view of files and folders. The 'hw03' folder contains 'rsc' and 'odt' subfolders. 'rsc' contains 'Configurations2', 'META-INF', 'Thumbnails', 'content.xml', 'manifest.rdf', 'meta.xml', 'mimetype', 'settings.xml', and 'styles.xml'. 'odt' contains 'jyun2zip', 'jyun.odt', 'jyun2edited.odt', 'jyun2edited2zip.zip', 'jyun2zip.zip', '2-01-odt-make.png', '2-02-odt-save.png', '2-03-hexeditor.png', '2-04-zip.png', '2-05-unzip.png', '2-06-mimetype.png', '2-07-wrap.png', '2-08-wrapped.png', '2-09-edit.png', '2-10-rezip.png', and '2-11-edited.png'. There are also 'hw03.md' and 'spec.pdf' files.
- Content Editor:** The main editor area shows the XML content of 'content.xml'. The XML is very large and complex, containing numerous namespace declarations (e.g., `xmlns:office="urn:openoffice:names:tc:opendocument:office:1.0"`, `xmlns:ooo="http://openoffice.org/2004/office"`, `xmlns:cals="urn:openoffice:names:experimental:calc:1.0"`) and detailed styling information for various document elements like tables, charts, and text blocks.

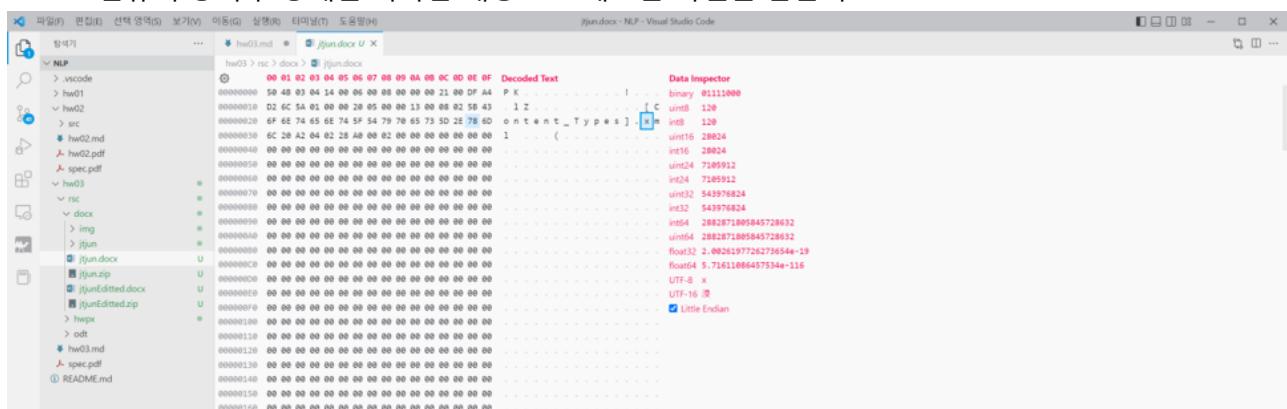
odt에서 새로운 라인을 추가하고 위의 과정을 반복해, content.xml 파일을 확인한 결과이다. 라인이 변경되었기 때문에 새로운 태그가 생기 것을 볼 수 있다

B. OOXML 중 DOCX 형식

1. MS Word로 DOCX 파일 만들기

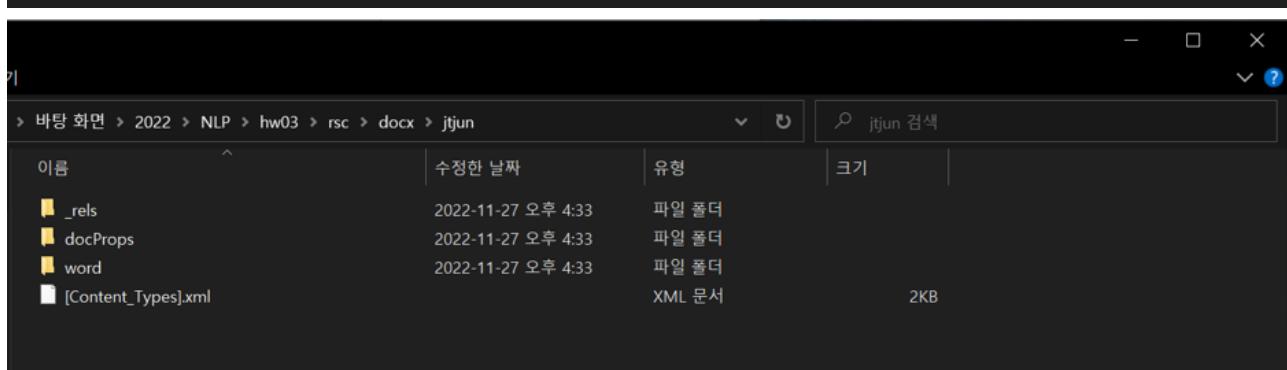
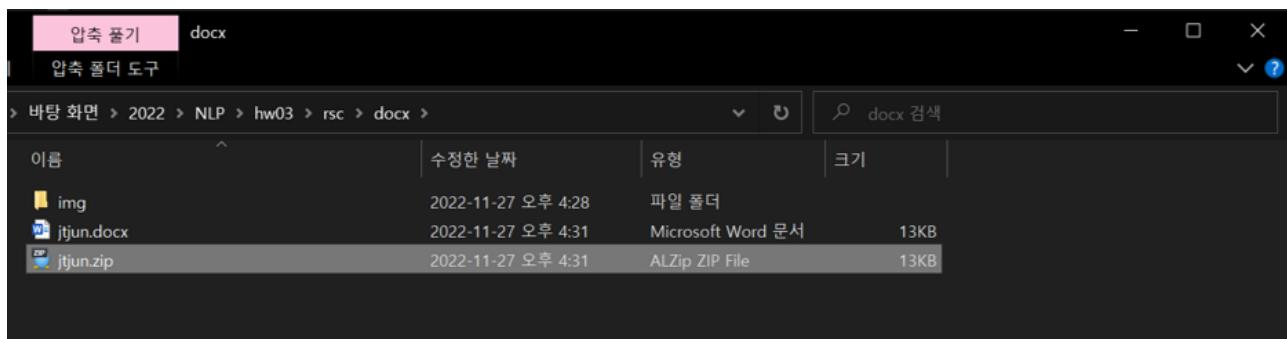


'DOCX 컴퓨터 공학부 장태준'이라는 내용으로 새로운 파일을 만든다.



VS Code에서 hex-editor로 docx파일을 확인한 결과는 위와 같다.

2. 압축 해제



확장자를 zip으로 변경한 뒤 압축해제 한다.

3. \word\document.xml 확인



The screenshot shows the Visual Studio Code interface with the following details:

- File Explorer:** On the left, it displays the project structure:
 - File (F)
 - Open (O)
 - 선택 영역 (V)
 - 보기 (V)
 - 이동 (G)
 - 설명 (R)
 - 터미널 (T)
 - 도움말 (H)
 - document.xml - NLP - Visual Studio Code
 - hw03.mrd
 - document.xml U
 - hw03.xml
 - hw03.docx
 - hw03.pdf
 - hw02.mrd
 - hw02.pdf
 - spec.pdf
 - hw02
 - src
 - rsc
 - docx
 - img
 - jrun
 - _rels
 - docProps
 - word
 - _rels
 - theme
 - document.xml
 - fontTable.xml
 - settings.xml
 - styles.xml
 - webSettings.xml
 - [Content_Types].xml
 - jrun.doc
 - jrun.zip
 - hwpx
 - ods
 - docx-01-create.png
 - docx-02-zip.png
 - docx-03-unzip.png
 - hw03.mrd
 - spec.pdf
 - README.md
- Editor:** The main editor area shows the XML code for `document.xml`. The code includes various namespaces such as `http://schemas.microsoft.com/office/word/2010/wordprocessingCanvas`, `http://schemas.microsoft.com/office/drawing/2015/10/21/charts`, and `http://schemas.microsoft.com/office/drawing/2016/5/12/charts`. It also contains styling elements like `w:color`, `w:font`, and `w:fontFamily`.

확장 기능을 통해 wrap 형식으로 `\word\document.xml` 파일을 확인해본다. 입력한 내용, '컴퓨터 공학부 장태준'이 들어가있는 것을 확인할 수 있다.

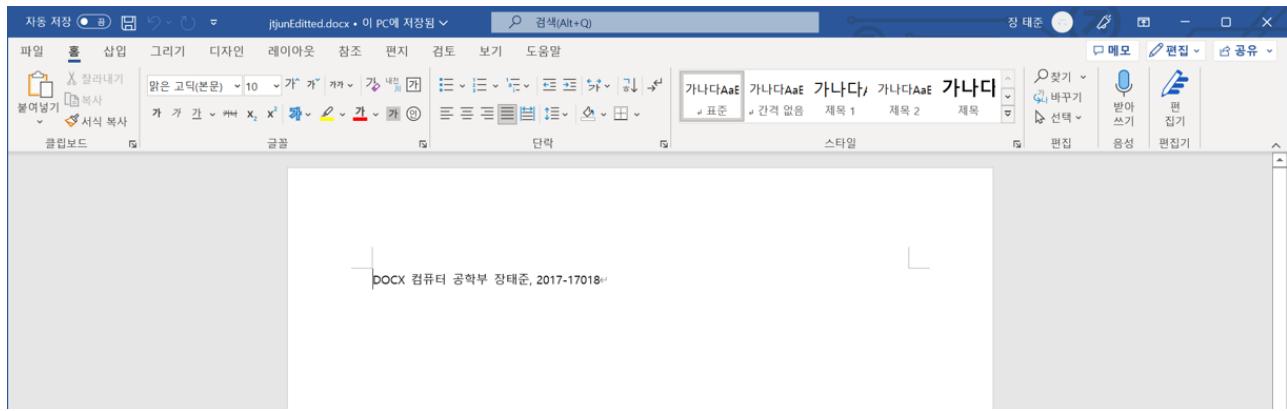
4. xml에서 변경 후 확인하기

기존 내용의 옆에 '2017-17018'를 추가한다

▶ 바탕 화면 > 2022 > NLP > hw03 > rsc > docx >					▼	↻	🔍 docx 검색
이름	▲	수정한 날짜	유형	크기	▶		
img		2022-11-27 오후 4:28	파일 폴더				
jjun		2022-11-27 오후 4:33	파일 폴더				
jjun.docx		2022-11-27 오후 4:31	Microsoft Word 문서	13KB			
jjun.zip		2022-11-27 오후 4:31	ALZip ZIP File	13KB			
jjunEdited.docx		2022-11-27 오후 4:36	Microsoft Word 문서	11KB			
jjunEdited.zip		2022-11-27 오후 4:36	ALZip ZIP File	11KB			

다시 압출한 뒤 확장자를 **docx**로 변경하다

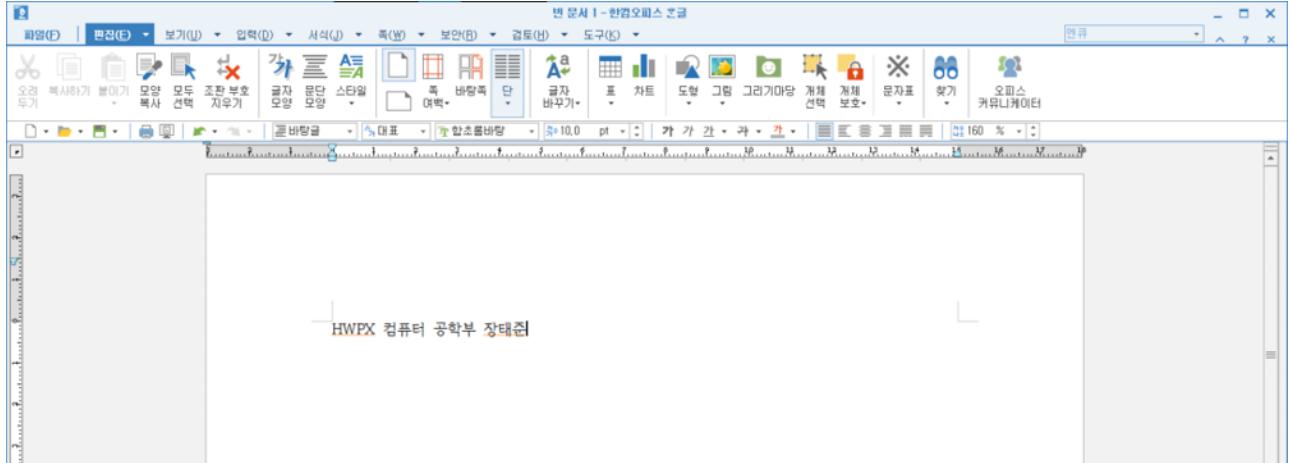
5. 결과



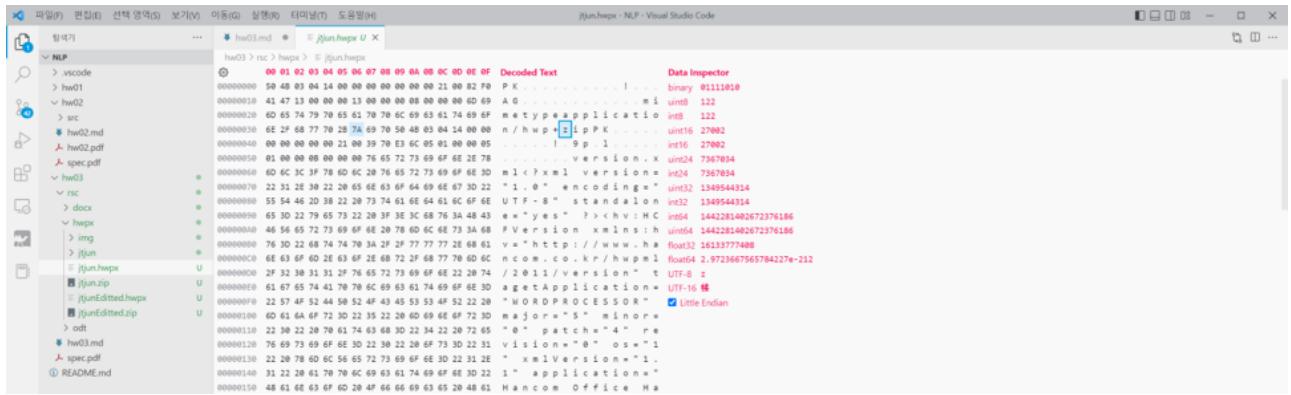
MS Word에서 위 파일을 확인한 결과, 변경한 내용을 확인할 수 있었다.

C. HWPX 형식

1. 한컴 오피스로 hwp 파일 만들기



'HWPX 컴퓨터 공학부 장태준'이라는 내용으로 새로운 파일을 만든다.



VS Code에서 hex-editor로 hpx파일을 확인한 결과는 위와 같다.

2. 압축 해제

Two screenshots of WinRAR showing the contents of the 'jtjun.zip' file. The top screenshot shows the contents of the 'hpx' folder, and the bottom screenshot shows the contents of the 'jtjun' folder.

파일	날짜	유형	크기
img	2022-11-27 오후 4:28	파일 폴더	
jtjun.hpx	2022-11-27 오후 4:39	한컴오피스 한글 표...	6KB
jtjun.zip	2022-11-27 오후 4:39	ALZip ZIP File	6KB

파일	날짜	유형	크기
Contents	2022-11-27 오후 4:40	파일 폴더	
META-INF	2022-11-27 오후 4:40	파일 폴더	
Preview	2022-11-27 오후 4:40	파일 폴더	
mimetype		파일	1KB
settings.xml		XML 문서	1KB
version.xml		XML 문서	1KB

확장자를 zip으로 변경한 뒤 압축해제 한다.

3. \Contents\section0.xml 확인

확장 기능을 통해 wrap 형식으로 `\Contents\section0.xml` 파일을 확인해본다. 입력한 내용, '컴퓨터 공학부 장태준'이 들어가있는 것을 확인할 수 있다.

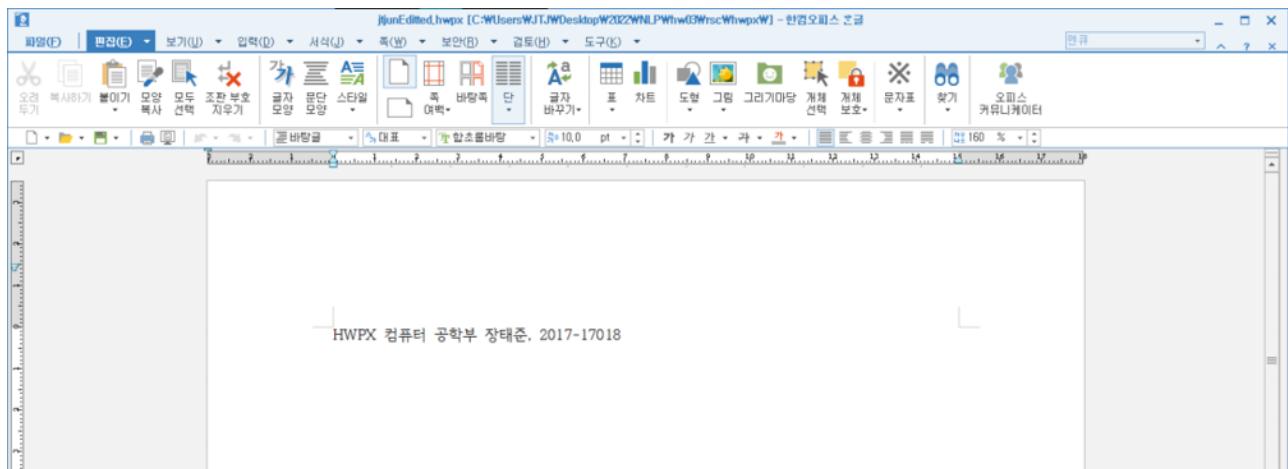
4. xml에서 변경 후 확인하기

기존 내용의 옆에 ', 2017-17018'를 추가한다.

파일 목록				
이름	수정한 날짜	유형	크기	操作
img	2022-11-27 오후 4:28	파일 폴더		
jtjun	2022-11-27 오후 4:40	파일 폴더		
jtjun.hwpx	2022-11-27 오후 4:39	한컴오피스 한글 표...	6KB	
jtjun.zip	2022-11-27 오후 4:39	ALZip ZIP File	6KB	
jtjunEdited.hwpx	2022-11-27 오후 4:44	한컴오피스 한글 표...	6KB	
jtjunEdited.zip	2022-11-27 오후 4:44	ALZip ZIP File	6KB	

다시 압축한 뒤, 확장자를 .hwpx로 변경한다.

5. 결과



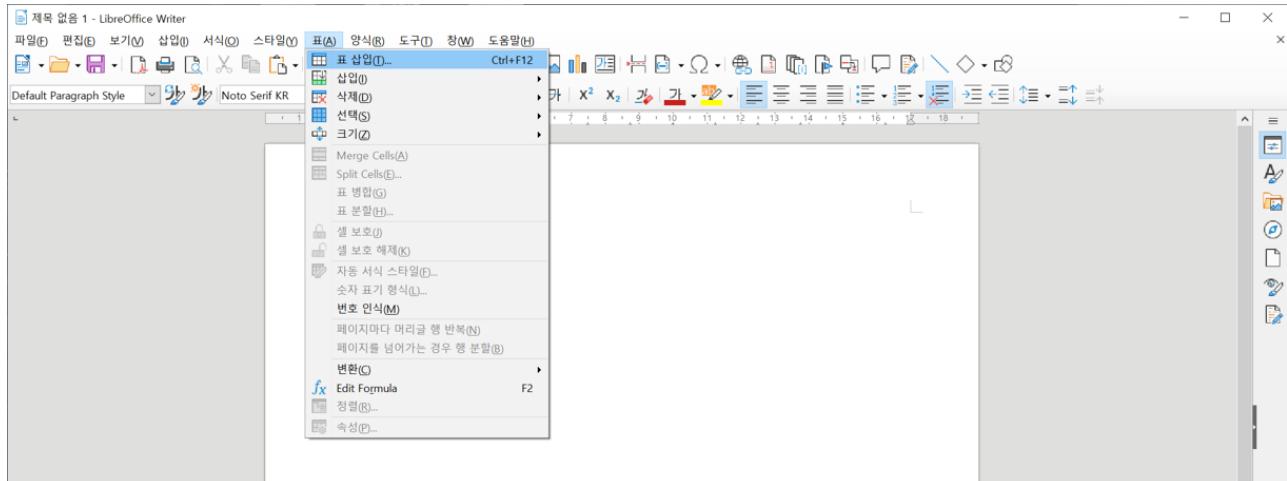
MS Word에서 위 파일을 확인한 결과, 변경한 내용을 확인할 수 있었다.

3. 테이블형 데이터

A. 간략한 테이블형 데이터 ODT 형식

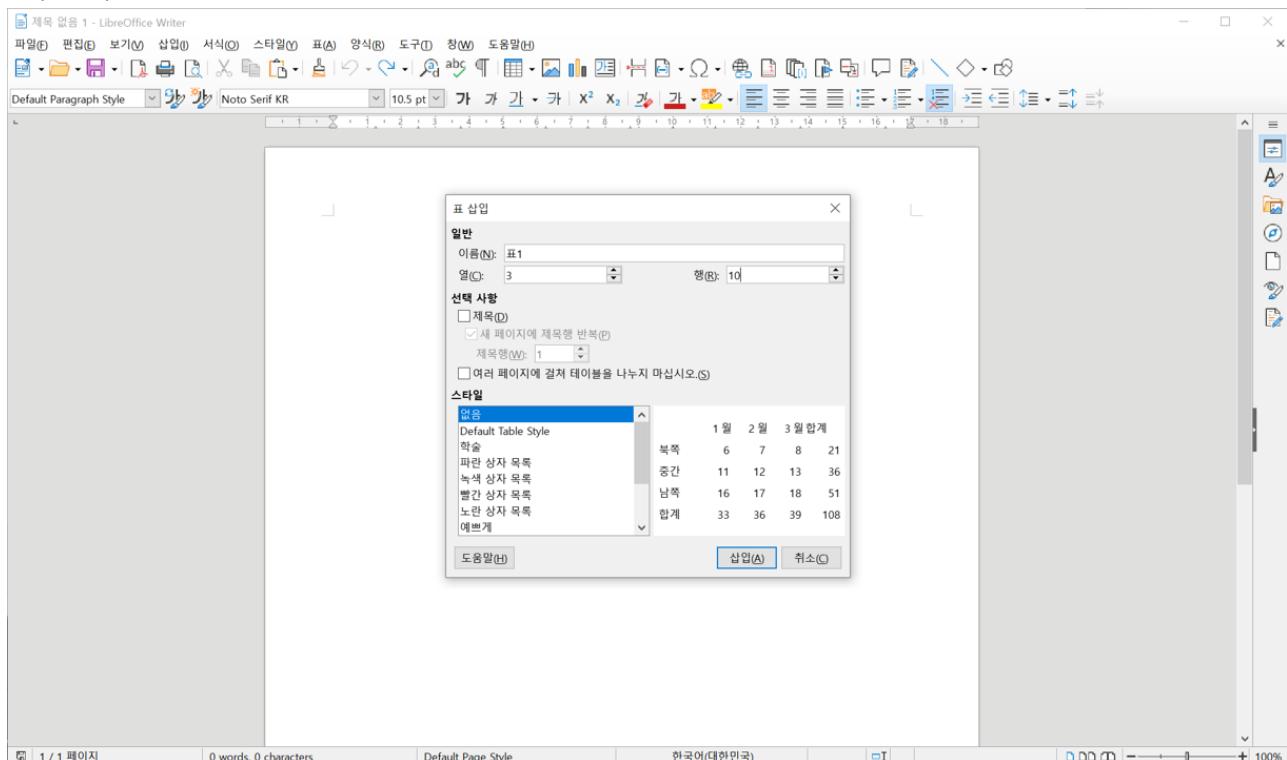
- 자유로운 주제로 만드시면 됩니다.
- 3열 이상 10행 이상의 규모

1. Libre Office에서 표 생성



상단의 도구표시줄에서 '표'를 클릭한 뒤, '표 삽입'을 클릭한다.

2. 표(table)의 크기 설정



행과 열을 설정하여, 표의 크기를 결정한다.

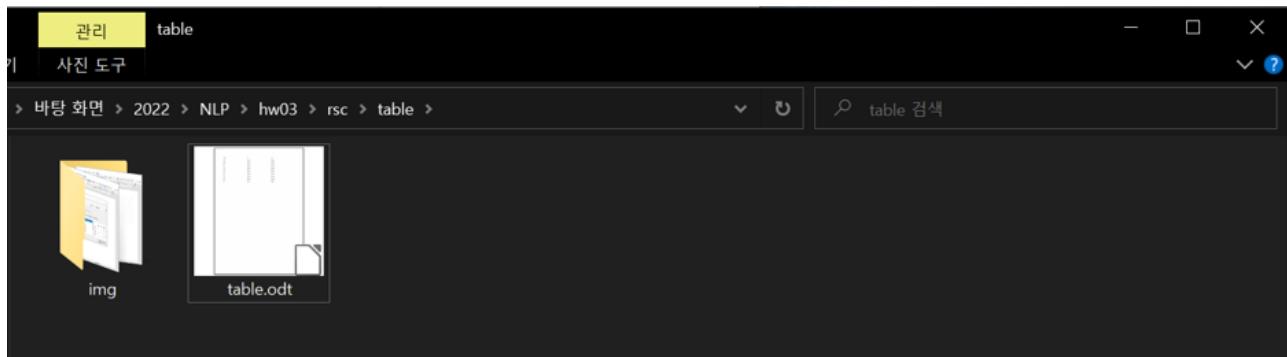
여기선, 과제의 요구사항에 맞춰, 3열 10행의 표를 생성하였다.

3. 내용 채우기

표	1	2
A	A-1	A-2
B	B-1	B-2
C	C-1	C-2
D	D-1	D-2
E	E-1	E-2
F	F-1	F-2
G	G-1	G-2
H	H-1	H-2
I	I-1	I-2

이후 실습에서 각 위치가 제대로 보존되는지 확인하기 위해 알파벳과 숫자를 조합에 좌표를 표시하였다. 행은 알파벳으로, 열은 숫자로 나타냈다.

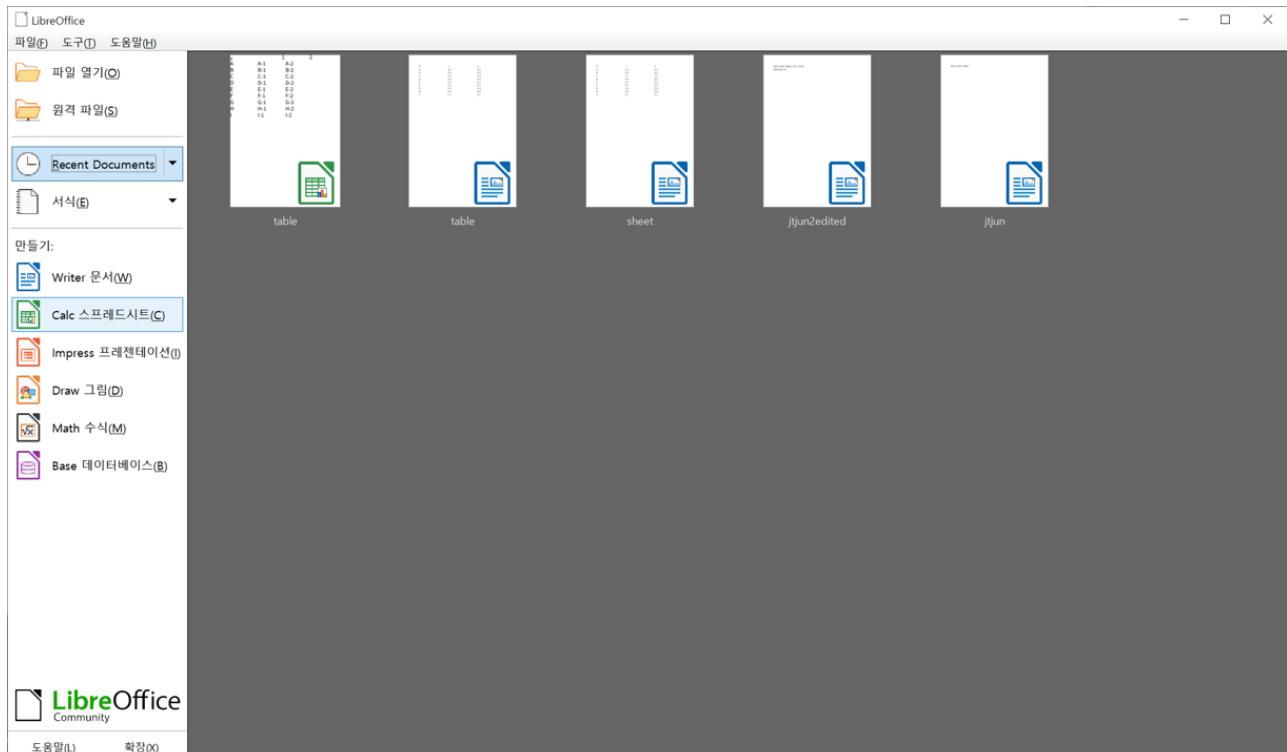
4. ODT 형식으로 저장



이렇게 생성한 표를 odt 형식으로 저장한다.

B. LibreOffice를 이용한 ODS 형식

1. Libre Office에서 표 생성



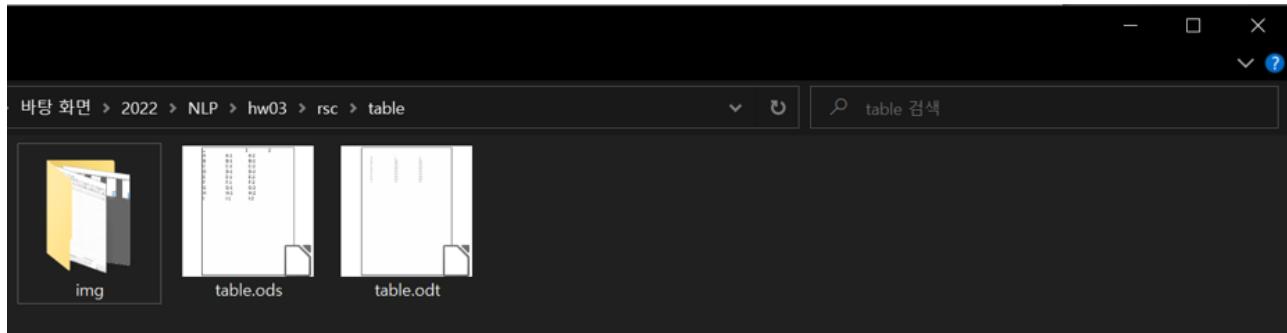
LibreOffice를 실행하여 calc 새 파일을 생성한다.

2. 내용 채우기

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	A	A-1	1	A-2												
2	B	B-1		B-2												
3	C	C-1		C-2												
4	D	D-1		D-2												
5	E	E-1		E-2												
6	F	F-1		F-2												
7	G	G-1		G-2												
8	H	H-1		H-2												
9	I	I-1		I-2												
10	I	I-1		I-2												
11																

A.에서 만든 표와 동일한 내용을 채운다.

3. ODS 형식으로 저장



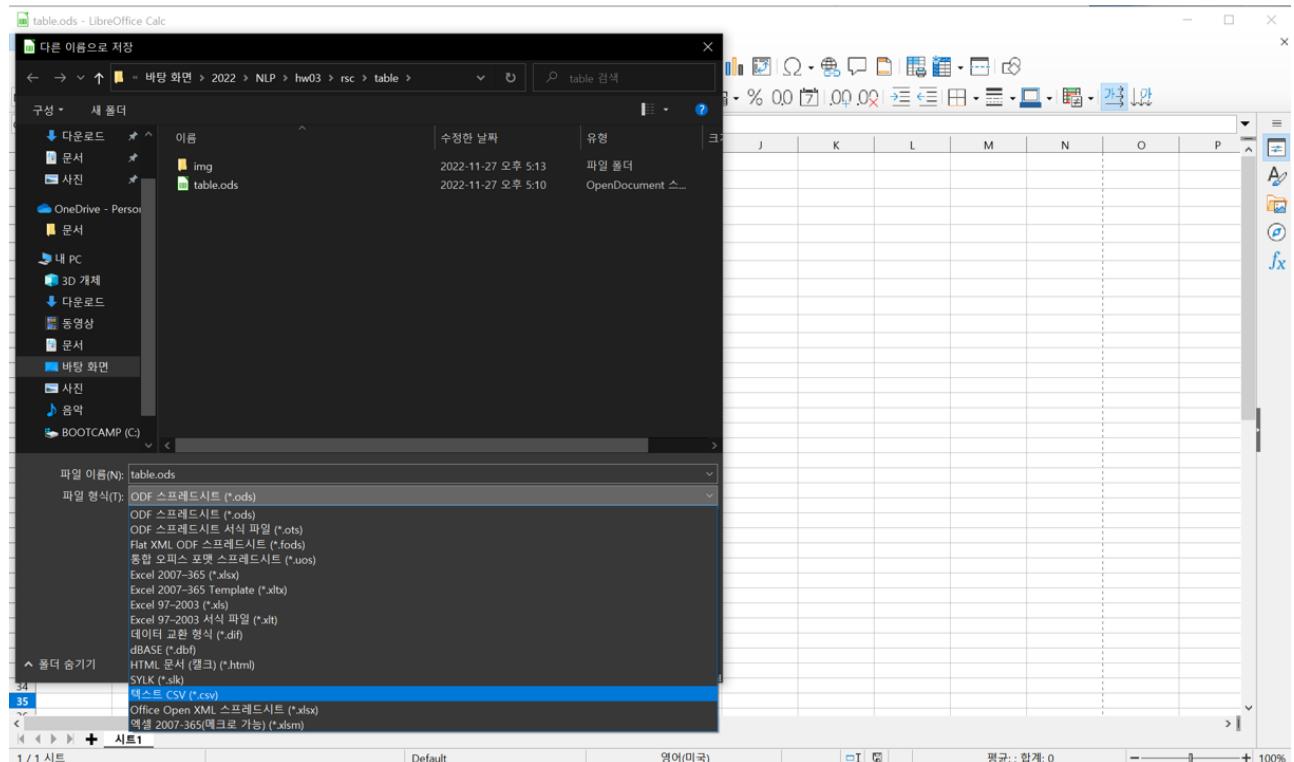
이렇게 생성한 표를 ods 형식으로 저장한다.

C. Visual Studio Code에서 CSV 형식

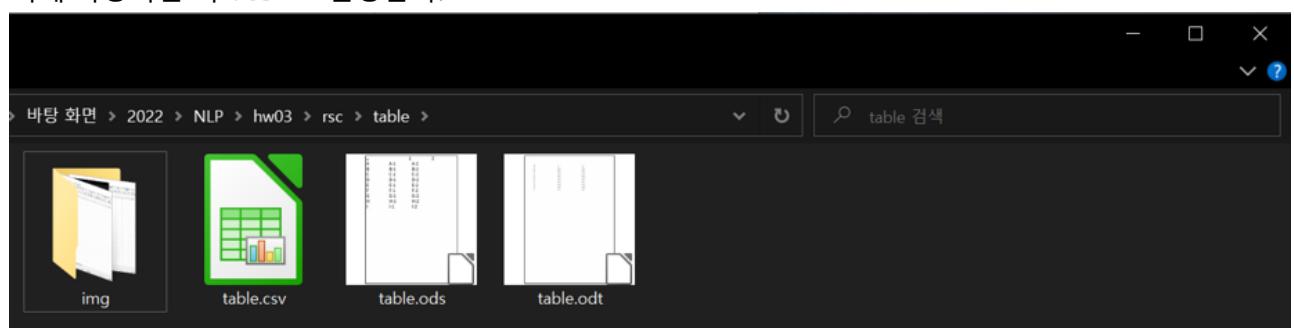
1. CSV 파일 만들기

파일(F)	편집(E)	보기(V)	삽입(I)	서식(O)	스타일(U)	시트(S)	데이터(D)	도구(U)	창(W)	도움말(H)
새로 만들기(N)										
열기(O)...	Ctrl+O									
원격 파일 열기(R)...										
최근 문서(U)										
닫기(C)										
마법사(W)										
서식 파일(M)										
다시 불러오기(L)										
버전(B)...										
저장(S)	Ctrl+S									
다른 이름으로 저장(A)...	Ctrl+Shift+S									
원격 저장(V)...										
복사본 저장(B)...										
모두 저장(N)										
내보내기(O)...										
PDF로 내보내기(E)...										
보내기(P)...										
웹 브라우저에서 미리보기(K)										
인쇄 미리보기(I)...	Ctrl+Shift+O									
인쇄(P)...	Ctrl+P									
프린터 설정(R)...										
속성(I)...										
디지털 서명(Q)...										
LibreOffice 종료(X)	Ctrl+Q									

B에서 만든 ODS 파일을 Libre Office로 연 후, 다른 이름으로 저장한다.

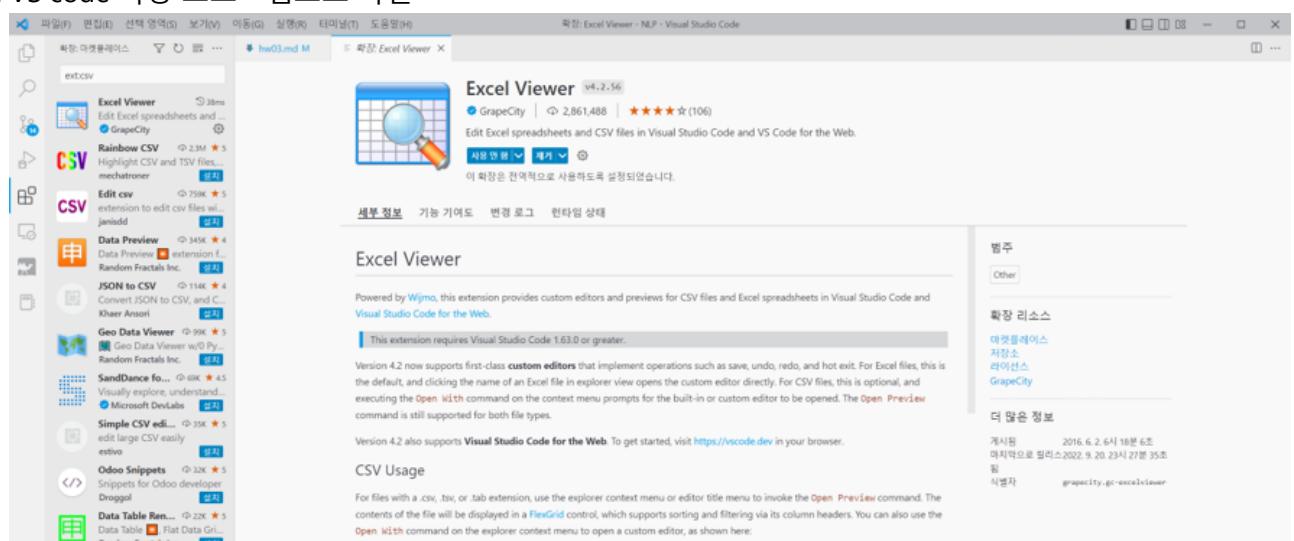


이때 확장자를 꼭 **.csv**로 설정한다.



CSV 파일이 생성된 것을 확인할 수 있다.

2. VS code 확장 프로그램으로 확인

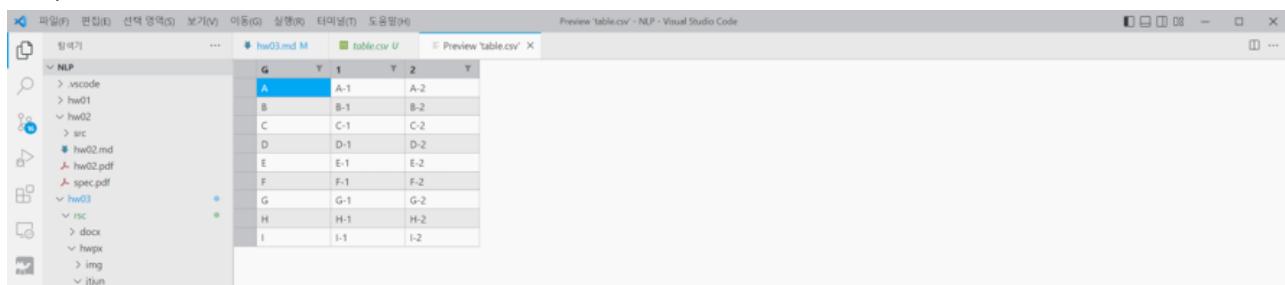


Excel Viewer라는 확장 프로그램을 설치한다.



CSV 파일을 연 뒤, 우측 상단에 있는 preview 버튼을 클릭한다.

3. 결과



첫번째 칸에 한글로 작성한 표가 깨져서 나오는 것을 확인할 수 있다.

D. LibreOffice를 이용한 TSV 형식

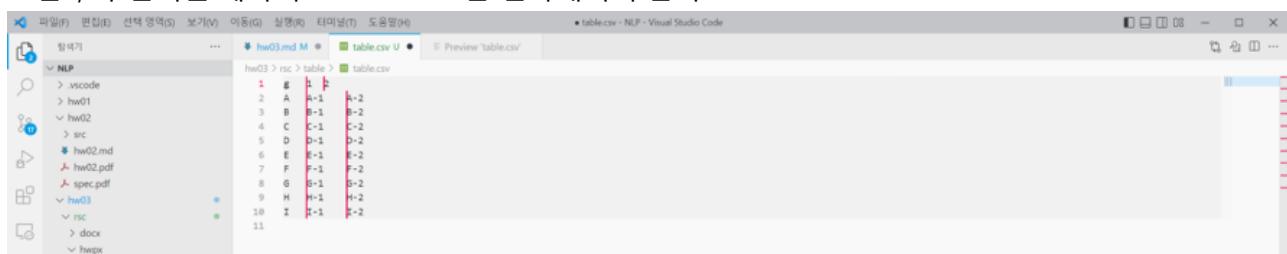
1. VS Code에서 CSV를 TSV로 바꾸기



VS Code에서 CSV 파일을 열고, 가장 처음에 있는 ','를 선택한다.



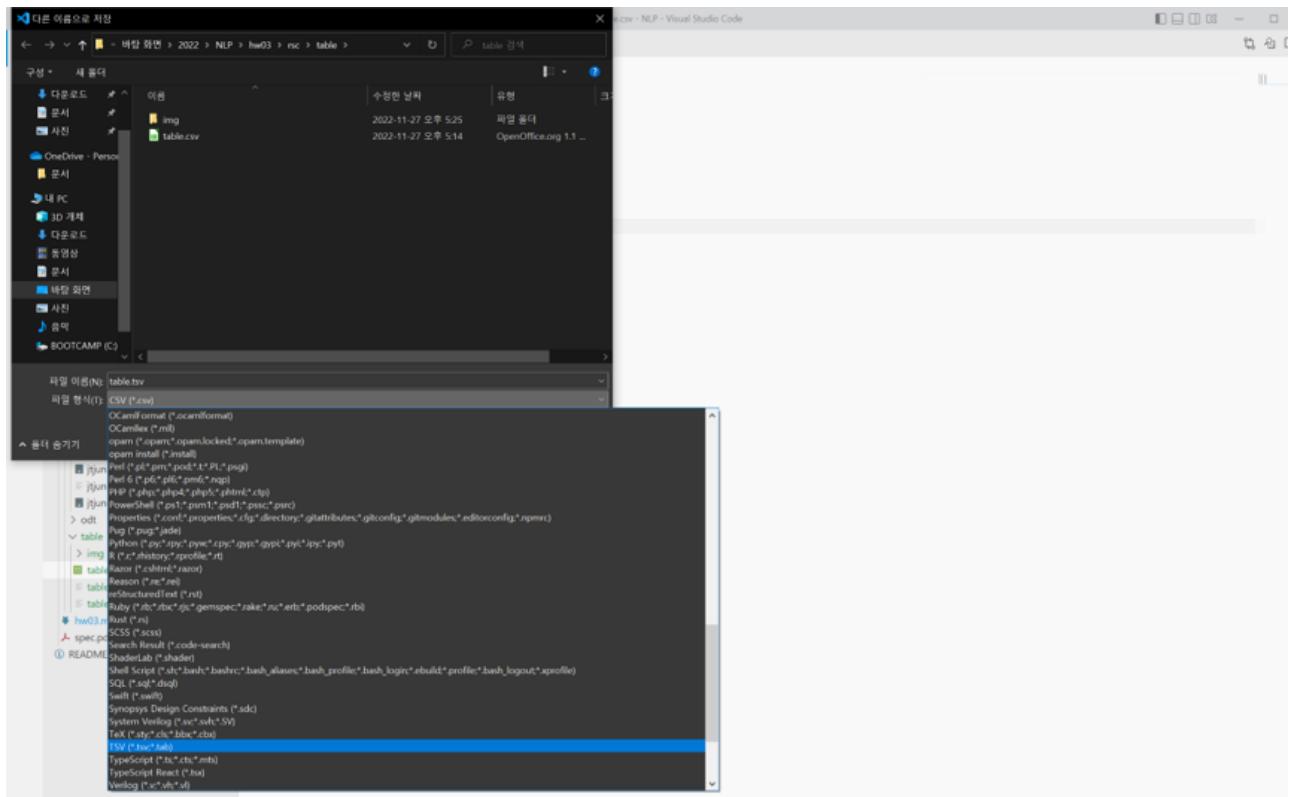
모든 ','가 선택될 때까지 **<Ctrl> + d**를 연속해서 누른다.



선택된 모든 ','를 **<Tab>**으로 변경한다.

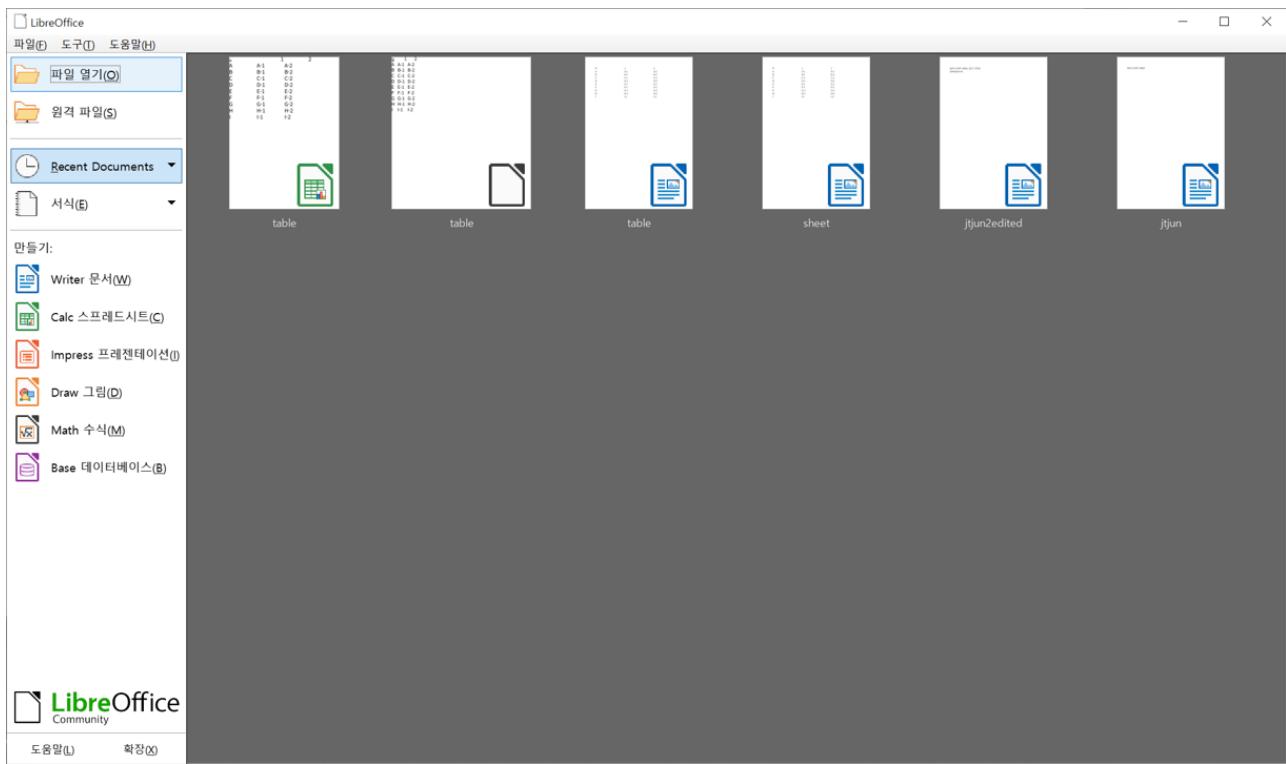
이때, VS Code의 설정에 따라 **<Tab>**이 아닌 여러 개의 **<Space>**로 변경될 수 있으니 주의한다.



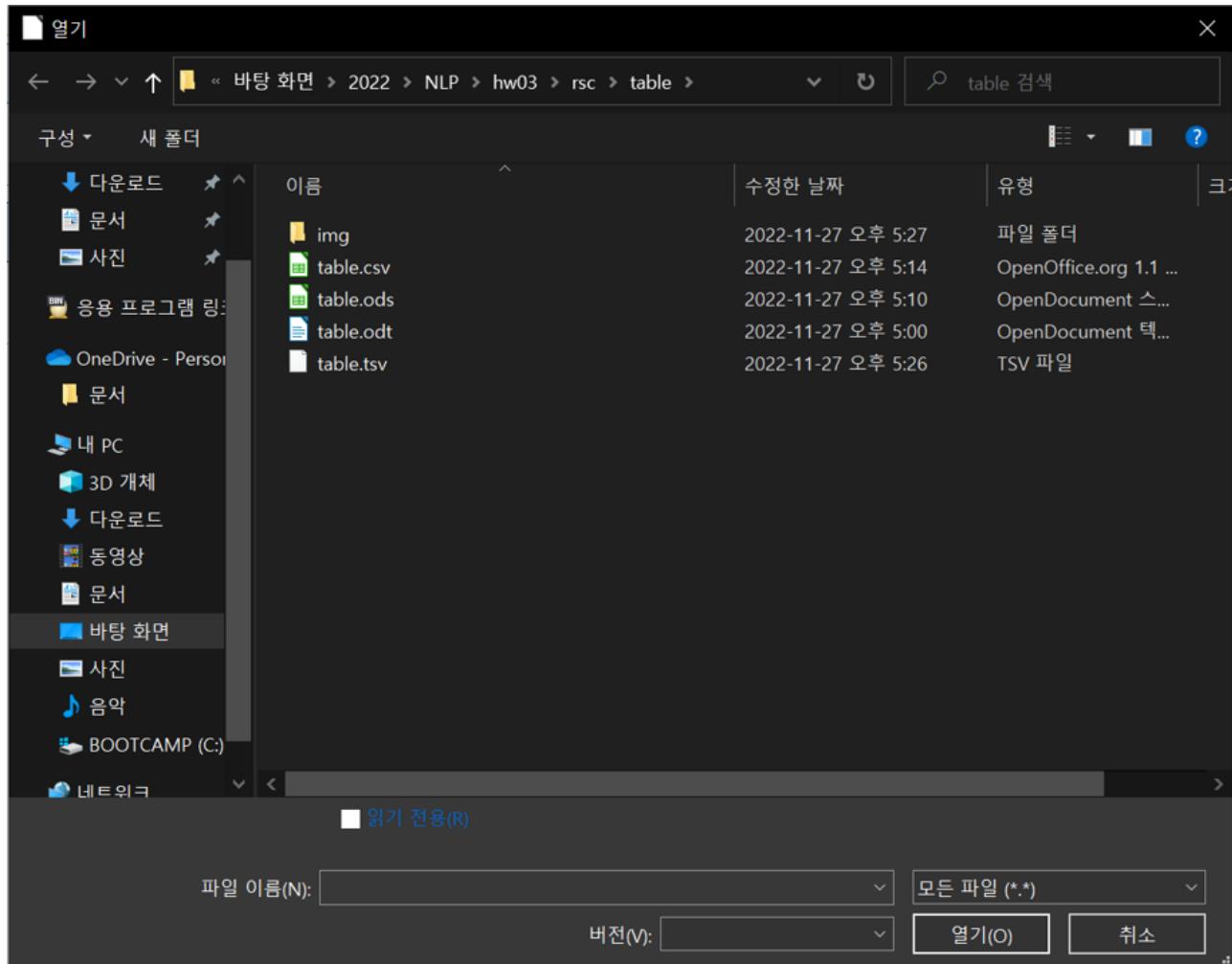


다른 이름으로 저장을 클릭하여, **.tsv** 확장자로 저장한다.

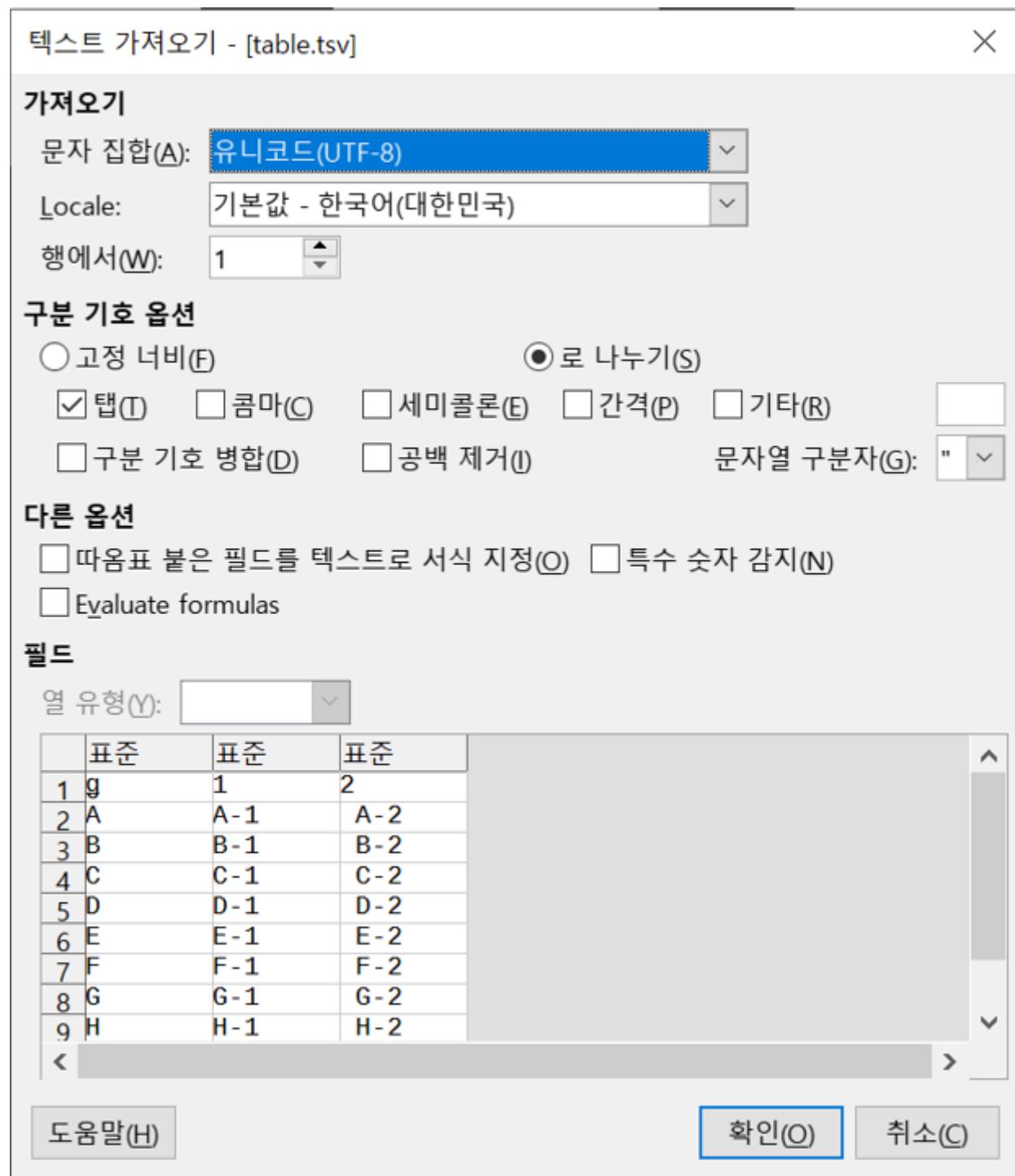
2. Libre Office에서 확인하기



Libre Office를 실행하여 '파일 열기' 버튼을 누른다.



3-D-1.에서 저장한 TSV 파일을 선택한다.



위

처럼 '텍스트 가져오기'가 나타나는데, 구분 기호 옵션을 '탭'으로 나누기로 설정한다.

3. 결과

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	g	1	2														
2	A	A-1	A-2														
3	B	B-1	B-2														
4	C	C-1	C-2														
5	D	D-1	D-2														
6	E	E-1	E-2														
7	F	F-1	F-2														
8	G	G-1	G-2														
9	H	H-1	H-2														
10	I	I-1	I-2														
11																	

CSV와 동일한 결과를 확인할 수 있다.

E. 소프트웨어들 사이에 데이터를 주고 받기

CSV

TSV

Excel

A	B	C	D	E	F	G
1		1	2			
2	A	A-1	A-2			
3	B	B-1	B-2			
4	C	C-1	C-2			
5	D	D-1	D-2			
6	E	E-1	E-2			
7	F	F-1	F-2			
8	G	G-1	G-2			
9	H	H-1	H-2			
10	I	I-1	I-2			
11						
12						
13						
14						
15						

깨쳤던 '표'글씨가 정상적으로 나타난다.

A	B	C	D	E	F	G
1		1	2			
2	A	A-1	A-2			
3	B	B-1	B-2			
4	C	C-1	C-2			
5	D	D-1	D-2			
6	E	E-1	E-2			
7	F	F-1	F-2			
8	G	G-1	G-2			
9	H	H-1	H-2			
10	I	I-1	I-2			
11						
12						
13						
14						
15						

Libre Office처럼 파일을 열 수 있었다.

Google

A	B	C	D	E	F	G
1		1	2			
2	A	A-1	A-2			
3	B	B-1	B-2			
4	C	C-1	C-2			
5	D	D-1	D-2			
6	E	E-1	E-2			
7	F	F-1	F-2			
8	G	G-1	G-2			
9	H	H-1	H-2			
10	I	I-1	I-2			
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						

'표'글씨가 깨졌다.

Naver

A	B	C	D	E	F	G
1		1	2			
2	A	A-1	A-2			
3	B	B-1	B-2			
4	C	C-1	C-2			
5	D	D-1	D-2			
6	E	E-1	E-2			
7	F	F-1	F-2			
8	G	G-1	G-2			
9	H	H-1	H-2			
10	I	I-1	I-2			
11						
12						
13						
14						
15						
16						
17						
18						

경고 메시지가 나타났으나, 정상적으로 파일을 열 수 있었다.

A	B	C	D	E	F	G
1		1	2			
2	A	A-1	A-2			
3	B	B-1	B-2			
4	C	C-1	C-2			
5	D	D-1	D-2			
6	E	E-1	E-2			
7	F	F-1	F-2			
8	G	G-1	G-2			
9	H	H-1	H-2			
10	I	I-1	I-2			
11						
12						
13						
14						
15						
16						
17						
18						

TSV를 지원하지 않아 파일이 열리지 않고 다운로드 되었다.