

Field research

This is subtitle

이진주

21 June, 2021

차 례

1	연구 개요	1
1.1	연구의 필요성	1
1.2	연구 목표	2
2	연구 방법	2
2.1	연구 방법 1 : Data collection_iGEM & Rmarkdown practice	2
2.2	연구 방법 2 : Rmarkdown practice & iGEM report	2
3	연구 결과	3
3.1	연구 결과 1 : github page 생성	3
3.2	연구 결과 2	4
4	고찰	4

1 연구 개요

R 프로그램을 이용하여 다양한 연구 결과를 종합, 정리하여 표준화할 수 있음.

합성생물학 분야의 대회인 iGEM 사례를 분석하며 R 프로그램을 이용해 데이터 처리, 분석을 할 수 있음.

1.1 연구의 필요성

R 프로그램을 이용하는 장점

생물 분야는 시스템의 복잡성 때문에 실험 재현이 힘들. 따라서 여러 연구자들의 반복 실험을 통해 특정 바이오파트의 표준화가 필요함.

합성생물학은 생물학에 공학적 개념을 적용함으로써 생물 시스템을 비교적 예측 가능하도록 만들기 위한 분야이다. 전통적인 생물 실험이 연역적인 실험과 관찰, 발견에 기반했다면 합성생물학에서는 생물 시스템을 설계함으로써

발명으로 전환한다고 볼 수 있음.

Rmarkdown의 철학과 필요성에 대한 참고영상

link : https://www.youtube.com/watch?v=s9aWmU0atIQ&ab_channel=RStudio

1.2 연구 목표

iGEM은 합성생물학 발전의 원동력이 되었다고 볼 수 있는데, 바이오파트를 표준화하거나 이를 이용한 유전자회로를 설계, 환경이나 보건 등의 목표를 가지고 연구하는 대회다.

해당 현장연구 수업에서는 합성생물학의 개념을 정립하고, 사용된 부품/회로들의 정량적 데이터를 수집하여 재현성을 분석한다. 그 과정에서 R 프로그램을 이용하여 Rmarkdown/Rstudio를 활용한다.

이곳에는 연구 목표를 적습니다.

2 연구 방법

2.1 연구 방법 1 : Data collection_iGEM & Rmarkdown practice

Rstudio는 R언어 외에도 다양한 언어를 이용한 프로그래밍을 지원하여 호환성이 좋으며, Rmarkdown, shiny 등을 활용하여 소통할 수 있음.

Rmarkdown에서 code chunk를 추가하여 코드를 작성 (단축키 Ctrl+Alt+i)

iGEM 홈페이지에서 5-10개 팀을 선정, 각 팀의 이름과 위키페이지 링크, 연구 내용 요약 등을 작성함.

- 팀이름
- 소속 조직
- 제목
- 분류
- wiki page
- 해결하고자 하는 문제
- 주요 해결 방법
- 사용한 부품

“igem team search_210329.pdf”

Knit 버튼에서 html, pdf, docx 중 원하는 포맷의 문서를 선택하면 해당 디렉토리에 파일이 생성됨.

2.2 연구 방법 2 : Rmarkdown practice & iGEM report

iGEM team 정리

실험에 사용한 방법, 사용한 DNA 부품과 회로를 파악하고 데이터화함.

R의 데이터는 vector로 처리되며, numeric, logical, character 등 크게 세 가지로 나눌 수 있음.

vector 값의 유형을 알 수 있는 함수 : class()

Combine function인 c()를 활용하여 vector 지정

다음은 값을 저장하고 그 값의 유형을 알 수 있는 R 코드와 지정한 데이터값으로 data frame을 만드는 코드임.

```
v1 <- c(1, 2, 3, 4)
v2 <- c("a", "b", "c", "d")
class(v1)
```

```
## [1] "numeric"
```

```
v_df <- data.frame(v1,v2)
v_df
```

```
##   v1 v2
## 1  1  a
## 2  2  b
## 3  3  c
## 4  4  d
```

3 연구 결과

3.1 연구 결과 1 : github page 생성

link: <https://jinjulee119.github.io/researcheweb/>

3.1.1 Rstudio에서 project 생성

- Rstudio > File > New Project > New Directory
- Directory name 입력 후 create project

3.1.2 Local project를 GitHub repository에 연결

- Rstudio > Tools > Version Control > Project Setup > Git/SVN
- Version control system에서 git 선택

3.1.3 Local 저장소에 commit

- Rstudio 상단 GIT 아이콘 > Commit (단축기 Ctrl+Alt+M)
- upload하고자 하는 파일 staged에 체크
- Commit message 입력 후 Commit 버튼 클릭

- 팝업창 close 후 우측 상단 Push 클릭, 창 닫기

3.2 연구 결과 2

3.2.1 iGEM_team table

[igem_team.xlsx]

이것은 연구 결과2 입니다. 이것은 연구 결과2 입니다. 이것은 연구 결과2 입니다. 이것은 연구 결과2 입니다.
 이것은 연구 결과2 입니다. 이것은 연구 결과2 입니다. 이것은 연구 결과2 입니다. 이것은 연구 결과2 입니다.
 이것은 연구 결과2 입니다.

4 고찰

이곳은 고찰을 적는 곳 입니다. 이곳은 고찰을 적는 곳 입니다. 이곳은 고찰을 적는 곳 입니다. 이곳은 고찰을 적는
 곳 입니다. 이곳은 고찰을 적는 곳 입니다. 이곳은 고찰을 적는 곳 입니다.

```
tmpdat2 <- igem_obs %>% full_join(igem_device, by=c("device_id"="id", "filename"="filename"))
%>% drop_na()
```

```
tmpdat2 %>% str
```

위 코드를 수행할 경우 drop_na() 코드로 인해 NA가 포함된 모든 row가 제거됨.

사용자에 따라 igem_obs의 promoter column이 없는 경우도 있고, strain이 비어있는 경우도 있기 때문에 위
 코드 그대로 수행할 경우 tmpdat2가 0 obs. of 15 variables, 15개의 column을 가지고 있지만 data가 없는
 table로 만들어짐.

따라서 NA가 없는 column을 지정해주면 해당 column에서 NA가 있는 row만 제거해야 함.

```
tmpdat2 <- igem_obs %>% full_join(igem_device, by=c("device_id"="id", "filename"="filename"))
```

위 코드를 수행할 경우 table of 146 obs. of 15 variables

```
tmpdat2 <- igem_obs %>% full_join(igem_device, by=c("device_id"="id", "filename"="filename"))
%>% select(id,strain,indc,conc,concunit,value,valunit,incubhr,incubtemp,device_id,device_name,part_combination,fi
%>% drop_na()
```

위 코드를 수행할 경우 table of 115 obs. of 13 variables 생성됨.