

Chapter 1. SAS 프로그램의 작성과 실행

SAS 시스템의 주요기능

❖ SAS(Statistical Analysis System)

- 다양한 유형의 자료를 포괄적으로 처리. 분석할 수 있도록 제반 수단을 제공하는 통계분석 시스템
- 다른 소프트웨어에 비해 시스템의 접근과 이용절차가 비교적 편리함
- 데이터 파일을 효율적으로 관리할 수 있음
- 다른 소프트웨어에 의해 만들어진 시스템 파일들을 직접 이용하는 것이 가능
- 기본적인 SAS 제품 이외에 다양한 수요에 호응하여 공급되는 여러 가지 제품들을 사용할 수 있음

SAS 시스템의 주요기능

❖ Base SAS

- SAS 시스템의 기본이 되는 제품
- 데이터셋의 입출력과 저장, 파일의 편집과 관리, 기본적인 통계분석 및 보고서 작성의 기능을 가짐
- 주요 기능
 - 데이터의 입력과 출력
 - 데이터의 수정과 프로그래밍
 - 정보의 출력 및 보고서 작성
 - 기본적인 통계자료 분석
 - 파일의 관리 및 조작

❖ SAS/STAT

- 전통적으로 SAS시스템에서 가장 많이 이용되는 제품
- 각종 통계분석의 처리를 포괄적으로 수용

❖ SAS Enterprise Guide

- 데이터의 접근 및 가공, 다양한 리포트 작성, 통계적인 분석 등의 과정이 프로그래밍 없이 메뉴 방식으로 가능하도록 지원
- 초보적인 SAS 사용자들에게 유용함

SAS 프로그램의 작성 및 결과 보기

SAS 프로그램의 구조

❖ 프로그램의 구성요소

- SAS 명령문(statement): 규정된 어휘와 이들을 전후 관련시키는 문법 체계에 의해 작성된 것으로, 특별한 명령을 수행하도록 지시하는 역할
 - 중심어(Keyword)
 - 하나의 SAS 명령문의 종류를 나타내는 것
 - Ex: DATA, INPUT, IF, CARDS, PROC, RUN
 - 이름(Name)
 - 변수, 데이터셋, 프로시저, 포맷 등의 이름을 나타냄
 - 변수와 데이터셋의 이름은 영문자 또는 _로 시작하는 이름을 사용자가 지정
 - 숫자의 사용은 가능하나, 공백, ., #, \$, ? 등의 특수문자는 사용할 수 없음
 - 특수문자와 연산자
 - 프로그램 내에서 특수한 역할을 하는 특수문자와 연산자를 사용할 수 있음
 - Ex: “\$”, “,”, “;”, “#”

SAS 프로그램의 구조

❖ 명령문의 형식

- 하나의 중심어로 시작되고 반드시 ;(semicolon)으로 닫아주어야 함
- 명령문의 각 원소들은 적어도 하나 이상의 공백(space)에 의해서 구분되어야 함
- 한 행의 어느 열에서나 시작할 수 있음
- 다음 행으로 한 단어의 짚림이 없이 연속될 수 있음
- 한 행에 여러 개의 명령문을 사용할 수 있음
- /* 주석(comment) */과 같은 형식으로 프로그램 내에 주석 추가 가능

명령문 예제

```
DATA ex1_1;  
INPUT id $ 1-3 name $ 5-10 sex $ 13 ewight 16-17 height 19-21;  
IF weight>=70 THEN weight1='GE70';  
ELSE weight1='LT70';  
CARDS;  
101 김철수 M 74 170  
102 이영희 F 68 166  
;  
RUN;
```

SAS 프로그램의 구조

❖ 데이터 단계와 프로시저 단계

- 데이터 단계(Data-step)
 - DATA라는 중심어와 데이터셋의 이름으로 시작
 - 새로운 데이터셋의 생성 및 자료값의 변형 등 데이터에 대한 조작을 수행하는 단계
 - 데이터셋의 이름은 규칙에 따라 사용자가 지정
- 프로시저 단계(Proc-step)
 - PROC이라는 중심어와 특정한 역할을 수행하는 프로시저의 이름으로 시작
 - 통계분석, 그래프 및 보고서 작성 등을 수행하는 단계
 - 프로시저 이름은 SAS에 등록된 이름을 사용
 - Ex: PRINT, PLOT, MEANS ...

데이터 단계와 프로시저 단계

```
DATA ex1_1;  
INPUT id $ 1-3 name $ 5-10 sex $ 13 eweight 16-17  
height 19-21;  
RUN;
```

데이터 단계

```
PROC MEANS DATA=ex1_1;  
VAR weight height;  
RUN;
```

프로시저 단계

연습문제

프로그램

```
DATA pilotdata;  
INFILE 'c:\temp\pilot.data';  
INPUT EmployeeID $ 1-6 FirstName $ 7-19 LastName $ 20-34 JobCode $ 35-  
41 Salary 42-47 Category $ 48-50;  
RUN;  
  
PROC SORT DATA=pilotdata OUT=sortedpilotdata;  
BY LastName FirstName;  
RUN;  
  
PROC PRINT DATA=sortedpilotdata;  
VAR EmployeeID FirstName LastName Jobcode;  
RUN;
```

- ❖ SAS 프로그램이 몇 단계로 이루어져 있는가?
- ❖ 데이터 단계는 몇 개인가?
- ❖ 프로시저 단계는 몇 개인가?
- ❖ 사용된 프로시저 이름은 무엇인가?
- ❖ 몇 개의 명령문으로 이루어져 있는가?